A Poisson INAR(1) process with a seasonal structure

4 authors:

Marcelo Bourguignon
Universidade Federal do Rio Grande do Norte

110 PUBLICATIONS   1,866 CITATIONS

SEE PROFILE

Klaus L.P. Vasconcellos
Federal University of Pernambuco

40 PUBLICATIONS   729 CITATIONS

SEE PROFILE

Valdério Reisen
Universidade Federal do Espírito Santo

23 PUBLICATIONS   451 CITATIONS

SEE PROFILE

Márton Ispány
University of Debrecen

60 PUBLICATIONS   623 CITATIONS

SEE PROFILE

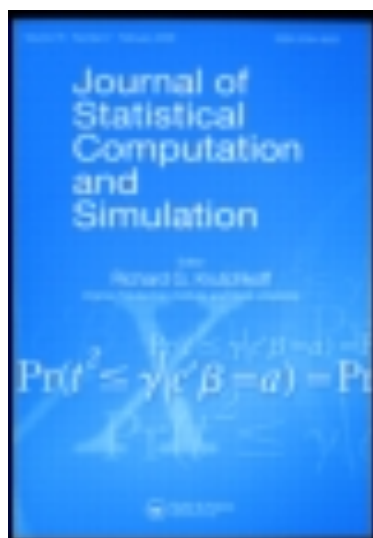**Some of the authors of this publication are also working on these related projects:**

Project   GAM-PCA-VAR View project

Project   INAR processes View project

Journal of Statistical Computation and Simulation

A Poisson INAR(1) process with a seasonal structure

Marcelo Bourguignon[a], Klaus L.P. Vasconcellos[a], Valdério A. Reisen[b] & Márton Ispány[c]

[a] Departamento de Estatística, Universidade Federal de Pernambuco, Recife, Brazil

[b] Departamento de Estatística and PPGEA, Universidade Federal do Espírito Santo, Vitória, Brazil

[c] Faculty of Informatics, University of Debrecen, Debrecen, Hungary
Published online: 19 Feb 2015.

Taylor & Francis
Taylor & Francis Group

A Poisson INAR(1) process with a seasonal structure

Marcelo Bourguignon[a], Klaus L.P. Vasconcellos[a*], Valdério A. Reisen[b] and Márton Ispány[c]

aDepartamento de Estatística, Universidade Federal de Pernambuco, Recife, Brazil; bDepartamento de Estatística and PPGEA, Universidade Federal do Espírito Santo, Vitória, Brazil; cFaculty of Informatics, University of Debrecen, Debrecen, Hungary

This paper introduces a non-negative integer-valued autoregressive (INAR) process with seasonal structure of first order, which is an extension of the standard INAR(1) model proposed by Al-Osh and Alzaid [First-order integer-valued autoregressive (INAR(1)) process. J Time Ser Anal. 1987;8:261–275]. The main properties of the model are derived such as its stationarity and autocorrelation function (ACF), among others. The conditional least squares and conditional maximum likelihood estimators of the model parameters are studied and their asymptotic properties are established. Some detailed discussion is dedicated to the case where the marginal distribution of the process is Poisson. A Monte Carlo experiment is conducted to evaluate and compare the performances of these estimators for finite sample sizes. The standard Yule–Walker approach is also considered for comparison purposes. The empirical results indicate that, in general, the conditional maximum likelihood estimator presents much better performance in terms of bias and mean square error. The model is illustrated using a real data set.

Keywords: INAR(1) model; conditional least squares; conditional maximum likelihood; seasonal period; Yule–Walker

1. Introduction

Seasonal time-series models have been extensively explored in the literature (see [1]). The smallest time period for this repetitive phenomenon to occur is called the seasonal period. Many business and economic time series display a seasonal phenomenon that repeats itself after a regular period of time. Seasonal phenomena may stem from factors such as weather, which affects many business and economic activities, prices of agricultural products, which have their supply usually related to the biological cycle of agricultural crops or livestock, cultural events closely related to sales and so on.

Over the last three decades, there has been a growing interest in modelling discrete-valued time-series models, that is, series taking values on a finite or countably infinite set. Models for count data have been widely used in several areas of study for various phenomena. A number of models for stationary processes with discrete marginal distributions have been proposed.[2] One of these models is the integer-valued autoregressive (INAR) process of Al-Osh and Alzaid,[3] which has similar structure and properties to the standard real-valued autoregressive models. The INAR model has been extensively studied in the literature. For example, see the survey by Weiß,[4] Jung et al. [5] and Freeland and McCabe.[6] Monteiro et al. [7] had proposed the

periodic INAR model of order one with period T driven by a periodic sequence of independent Poisson-distributed random variables.

However, the study of seasonal time series of counts has not received much attention so far in the literature. This paper aims to give a contribution in this direction. The motivation for such a process arises from its potential in modelling and analysing a stationary non-negative integer-valued time series with seasonal patterns. Therefore, the study of seasonal extensions of the INAR processes motivates a novel research branch with many practical applications, such as modelling the number of hospital emergency service arrivals caused by diseases that present seasonal behaviour,[8] the monthly number of claims of short-term disability benefits made by injured workers in industry,[9] the number of traffic accidents,[10] the demand for travelling, highest during the warmer summer months and lowest during the winter months, the number of crime occurrences [11] and so on.

The Poisson INAR(1) model proposed by Al-Osh and Alzaid [3] is not suitable for modelling counting series that present seasonality. Zhu and Joe [12] generalized the INAR(1) model with Poisson and negative binomial innovations, respectively, by using covariates in the parameters of the innovations to explain seasonality. However, these models are not stationary. In this context, the objective of this paper is to propose a second-order stationary integer-valued model with seasonal patterns based on the model of Al-Osh and Alzaid.[3] Our main goal is to investigate basic probabilistic and statistical properties of the model presented here, as well as inferential methods for the parameters associated with it.

The paper is structured as follows. The model is formally defined in Section 2 and some of its properties are outlined. In Section 3, estimation methods for the model parameters are proposed. Section 4 discusses theoretical results in order to obtain point forecasts. Section 5 discusses some simulation results for the estimation methods. In Section 6, the model is applied to a well-known data set. Remarks and conclusions are stated in Section 7. Finally, the proofs of all propositions and theorems are contained in the appendix.

2. The first-order seasonal non-negative INAR model

Let $\mathbb{Z}_+$, $\mathbb{N}$ and $\mathbb{R}_+$ denote the set of non-negative integers, positive integers and non-negative real numbers, respectively. All random variables will be defined on a common probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

Let X be a non-negative integer-valued random variable and let $\phi \in [0, 1]$. According to Steutel and Van Harn,[13] the '∘' binomial thinning operator is defined as follows:

$$\phi \circ X = \sum_{j=1}^{X} Z_j,$$

where $\{Z_j\}_{j\in\mathbb{N}}$ are independent and identically distributed (i.i.d.) random variables, mutually independent of X, with $\mathbb{P}(Z_j = 1) = 1 - \mathbb{P}(Z_j = 0) = \phi$, that is, $\{Z_j\}_{j\in\mathbb{N}}$ is an i.i.d. Bernoulli random sequence with mean ϕ. Note that for $X = 0$, the empty sum is defined as 0. With this operator, the first-order seasonal non-negative INAR model is defined as follows.

DEFINITION 1 *A discrete-time non-negative integer-valued stochastic process* $\{Y_t\}_{t\in\mathbb{Z}}$ *is said to be a first-order seasonal INAR process with seasonal period s* (INAR(1)$_s$) *if it satisfies the following equation:*

$$Y_t = \phi \circ Y_{t-s} + \epsilon_t, \quad t \in \mathbb{Z}, \tag{1}$$

where $\phi \in [0, 1]$, $\{\epsilon_t\}_{t\in\mathbb{Z}}$ *is an innovation sequence of i.i.d non-negative integer-valued random variables not depending on past values of* $\{Y_t\}_{t\in\mathbb{Z}}$ *and* $s \in \mathbb{N}$ *denotes the seasonal period. It*

is also assumed that the Bernoulli variables that define $\phi \circ Y_{t-s}$, that is, the Bernoulli variables from which Y_t are obtained, are independent of the Bernoulli variables from which other values of the series are calculated. Moreover, we assume that all Bernoulli variables defining the thinning operations are independent of the innovation sequence $\{\epsilon_t\}_{t \in \mathbb{Z}}$.

The process $\{Y_t\}_{t \in \mathbb{Z}}$ in Equation (1) has two random components: the survivors of the elements of the process at time $t - s$, given by $\phi \circ Y_{t-s}$, each with probability ϕ of survival, and the elements which entered the system in the interval $(t - s, t]$, defining the innovation term ϵ_t. Note that the INAR(1) model [3] is a particular case of the INAR(1)$_s$ when $s = 1$. On the other hand, the INAR(1)$_s$ process consists of s mutually independent INAR(1) processes equipped with the same autoregressive coefficient ϕ and same innovation distribution. Namely define the processes $\{Y_t^{(r)}\}_{t \in \mathbb{Z}}, r = 0, 1, \ldots, s - 1$, as $Y_t^{(r)} := Y_{ts+r}, t \in \mathbb{Z}$. Then, it is easy to see by Equation (1) that, for all $r = 0, 1, \ldots, s - 1$, the process $\{Y_t^{(r)}\}_{t \in \mathbb{Z}}$ satisfies the INAR(1) model

$$Y_t^{(r)} = \phi \circ Y_{t-1}^{(r)} + \epsilon_t^{(r)}, \quad t \in \mathbb{Z},$$

where the innovation sequence $\{\epsilon_t^{(r)}\}_{t \in \mathbb{Z}}$ is defined by $\epsilon_t^{(r)} := \epsilon_{ts+r}, t \in \mathbb{Z}$. The independence of the stochastic processes $\{Y_t^{(r)}\}_{t \in \mathbb{Z}}, r = 0, 1, \ldots, s - 1$, clearly follows from the independences of the innovation sequences $\{\epsilon_t^{(r)}\}_{t \in \mathbb{Z}}, r = 0, 1, \ldots, s - 1$, and of the counting processes involved in the thinning operators. The stochastic processes $\{Y_t^{(r)}\}_{t \in \mathbb{Z}}, r = 0, 1, \ldots, s - 1$, are referred to as the seasonal components of the INAR(1)$_s$ process $\{Y_t\}_{t \in \mathbb{Z}}$. This decomposition implies that an INAR(1)$_s$ process $\{Y_t\}_{t \in \mathbb{Z}}$ is a so-called s-step Markov chain, that is, for all $t \geq s$,

$$\mathbb{P}(Y_t = y_t | Y_{t-1} = y_{t-1}, \ldots, Y_0 = y_0) = \mathbb{P}(Y_t = y_t | Y_{t-s} = y_{t-s}), \tag{2}$$

for any $y_0, y_1, \ldots, y_t \in \mathbb{Z}_+$. In addition, if $\phi = 1$, Equation (1) becomes $Y_t = Y_{t-s} + \epsilon_t, t \in \mathbb{Z}$, defining, in this case, a seasonal unit root process. However, due to the non-stationary behaviour of this process, it is not considered in this paper.

PROPOSITION 1 *If $\phi \in [0, 1)$, the unique stationary marginal distribution of model* (1) *can be expressed in terms of the innovation process $\{\epsilon_t\}_{t \in \mathbb{Z}}$ as*

$$Y_t \overset{d}{=} \sum_{k=0}^{\infty} \phi^k \circ \epsilon_{t-ks} = \epsilon_t + \sum_{k=1}^{\infty} \sum_{j=1}^{\epsilon_{t-sk}} Z_{t,k,j}, \quad t \in \mathbb{Z}, \tag{3}$$

where $\overset{d}{=}$ stands for equality in distribution. For all $t \in \mathbb{Z}$, the infinite series is understood as the limit in probability of the finite sum, the Bernoulli variables $\{Z_{t,k,j}\}_{k,j \in \mathbb{N}}$ being mutually independent and independent of the innovation process, with $\mathbb{E}Z_{t,k,j} = \phi^k$ for all $k, j \in \mathbb{N}$.

Proof See the appendix. ∎

Remark 1 Proposition 1 shows that the stationarity condition of the INAR(1)$_s$ process is equivalent to that of the INAR(1) process, that is, if $\phi \in [0, 1)$, the first-order seasonal INAR process has a unique stationary solution.

In this paper, it is assumed that $\{\epsilon_t\}_{t \in \mathbb{Z}}$ is an i.i.d. sequence of Poisson-distributed variables with mean $\lambda \in \mathbb{R}_+$ and that, for all t, this sequence is mutually independent of all Bernoulli random variables that define $\phi \circ Y_t$. Also, we denote by $\mathcal{P}o(\lambda)$, with $\lambda \in \mathbb{R}_+$, and by $\mathcal{B}(n, p)$, with $n \in \mathbb{Z}_+$ and $p \in [0, 1]$, the Poisson and the binomial distributions, respectively.
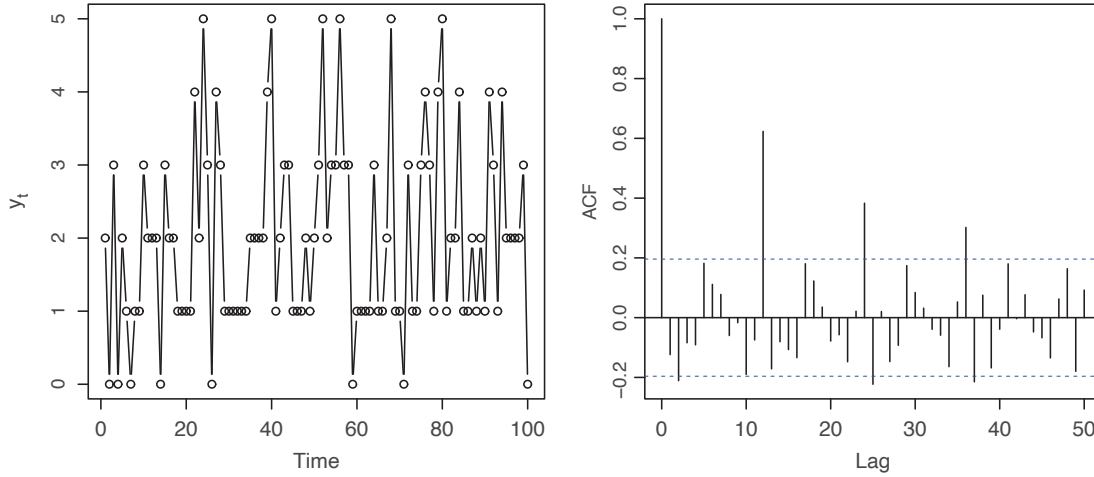
Figure 1. Sample path of INAR(1)$_s$ and its sample ACF.

The following proposition formalizes some properties of the INAR(1)$_s$ model.

PROPOSITION 2 *Let* $\{\epsilon_t\}_{t \in \mathbb{Z}}$ *be an i.i.d. sequence of Poisson-distributed variables with mean* $\lambda \in \mathbb{R}_+$ *and let* $\phi \in [0, 1)$. *Let* $\{Y_t\}_{t \in \mathbb{Z}}$ *be a stationary solution of the INAR(1)$_s$ model defined by Equation* (1). *Then,*

(a) Y_t*has Poisson distribution with parameter* $\lambda/(1 - \phi)$, *that is,* $Y_t \sim \mathcal{P}o(\lambda/(1 - \phi))$ *for all* $t \in \mathbb{Z}$ *and thus*

$$\mathbb{E}(Y_t) = \text{var}(Y_t) = \frac{\lambda}{1 - \phi}, \quad t \in \mathbb{Z}.$$

(b) *the ACF is given by*

$$\rho(k) = \begin{cases} \phi^{k/s} & \text{if } k \text{ is a multiple of } s, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Proof See the appendix. ∎

Remark 2 Proposition 2 states that if the innovation sequence is $\mathcal{P}o(\lambda)$ distributed, then $\mathcal{P}o(\lambda/(1 - \phi))$ provides the only stationary marginal distribution for Y_t. In this case, the process $\{Y_t\}_{t \in \mathbb{Z}}$ given in Definition 1 is called a Poisson INAR(1)$_s$ process.

Equation (4) shows that the ACF, $\rho(k)$, decays exponentially with lag k. Furthermore, the ACF of an INAR(1)$_s$ model is zero except at lags that are multiples of s. Figure 1 presents 100 simulated values of the INAR(1)$_s$ process and its sample ACF for $\phi = 0.5, \lambda = 1$ and $s = 12$.

3. Estimation methods

In what follows, some analytical and asymptotic results for the conditional least squares (CLS) and the conditional maximum likelihood (CML) estimators of the vector of parameters $\theta = (\phi, \lambda)^\mathsf{T}$, for the model defined in Equation (1), are derived. Let $Y_0, Y_1, \ldots, Y_n$ be a sample of a Poisson INAR(1)$_s$ process. For all $k \in \mathbb{Z}_+$, $\mathcal{F}_k$ denotes the σ-algebra generated by the random variables $Y_0, Y_1, \ldots, Y_k$. $\mathbb{P}_{\theta}(\cdot)$ and $\mathbb{E}_{\theta}(\cdot)$ denote the probability and the expectation with respect to the true parameter θ, respectively.

The CLS estimator $\hat{\boldsymbol{\theta}}_{\text{CLS}} = (\hat{\phi}_{\text{CLS}}, \hat{\lambda}_{\text{CLS}})^{\mathsf{T}}$ of $\boldsymbol{\theta} = (\phi, \lambda)^{\mathsf{T}}$ is given by

$$\hat{\boldsymbol{\theta}}_{\text{CLS}} := \arg\min_{\boldsymbol{\theta}}(Q_n(\boldsymbol{\theta})),$$

where $Q_n(\boldsymbol{\theta}) = \sum_{t=s+1}^{n}[Y_t - \mathbb{E}_{\boldsymbol{\theta}}(Y_t|\mathcal{F}_{t-1})]^2$. Since the INAR(1)$_s$ process $\{Y_t\}_{t \in \mathbb{Z}}$ satisfies Equation (1), we can easily obtain $\mathbb{E}_{\boldsymbol{\theta}}(Y_t|\mathcal{F}_{t-1}) = \mathbb{E}_{\boldsymbol{\theta}}(Y_t|Y_{t-s}) = g(\boldsymbol{\theta}, Y_{t-s})$, where the function $g(\cdot, \cdot)$ is defined as $g(\boldsymbol{\theta}, y) := \phi y + \lambda, y \in \mathbb{Z}_+$. Thus, the CLS estimates of ϕ and λ will be given, respectively, as

$$\hat{\phi}_{\text{CLS}} := \frac{(n-s)\sum_{t=s+1}^{n} Y_t Y_{t-s} - \sum_{t=s+1}^{n} Y_t \sum_{t=s+1}^{n} Y_{t-s}}{(n-s)\sum_{t=s+1}^{n} Y_{t-s}^2 - \left(\sum_{t=s+1}^{n} Y_{t-s}\right)^2} \tag{5}$$

and

$$\hat{\lambda}_{\text{CLS}} := \frac{1}{n-s}\left(\sum_{t=s+1}^{n} Y_t - \hat{\phi}_{\text{CLS}} \sum_{t=s+1}^{n} Y_{t-s}\right). \tag{6}$$

It can be verified that the function $g(\cdot, \cdot)$ together with its first and second partial derivatives with respect to its first argument satisfy all of the regularity conditions of Theorem 3.2 in [14]; see, also, the proof in the appendix. Therefore, the CLS estimators in Equations (5) and (6) are strongly consistent and asymptotically normally distributed. These are formalized in the next theorem.

THEOREM 3 *The CLS estimator* $\hat{\boldsymbol{\theta}}_{\text{CLS}} = (\hat{\phi}_{\text{CLS}}, \hat{\lambda}_{\text{CLS}})^{\mathsf{T}}$ *of the parameter* $\boldsymbol{\theta} = (\phi, \lambda)^{\mathsf{T}}$ *of a Poisson* INAR(1)$_s$ *process has the following asymptotic distribution*

$$\sqrt{n}\begin{pmatrix} \hat{\phi}_{\text{CLS}} - \phi \\ \hat{\lambda}_{\text{CLS}} - \lambda \end{pmatrix} \xrightarrow{d} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \lambda^{-1}\phi(1-\phi)^2 + (1-\phi^2) & -(1+\phi)\lambda \\ -(1+\phi)\lambda & \lambda + (1+\phi)(1-\phi)^{-1}\lambda^2 \end{pmatrix}\right).$$

Proof See the appendix. ∎

Following the standard notation, the conditional distribution of Y_t given Y_{t-s} is here written as $\mathbb{P}_{\boldsymbol{\theta}}(Y_t|Y_{t-s})$, which is defined by the conditional probabilities $\mathbb{P}_{\boldsymbol{\theta}}(Y_t = y_t|Y_{t-s} = y_{t-s}), y_t, y_{t-s} \in \mathbb{Z}_+$. The CML estimator is based upon the fact that the conditional distribution $\mathbb{P}_{\boldsymbol{\theta}}(Y_t|Y_{t-s})$ is the convolution of the binomial distribution $\mathcal{B}(Y_{t-s}, \phi)$ resulting from the random variable $\phi \circ Y_{t-s}$ with the Poisson distribution $\mathcal{P}o(\lambda)$ of the innovation ϵ_t. Then,

$$\mathbb{P}_{\boldsymbol{\theta}}(Y_t|Y_{t-s}) = [\mathcal{B}i(Y_{t-s}, \phi) * \mathcal{P}o(\lambda)](Y_t) = e^{-\lambda} \sum_{i=0}^{\min(Y_t, Y_{t-s})} \frac{\lambda^{Y_t-i}}{(Y_t-i)!}\binom{Y_{t-s}}{i}\phi^i(1-\phi)^{Y_{t-s}-i}.$$

Using property (2), the conditional likelihood function of the INAR(1)$_s$ process with true parameter $\boldsymbol{\theta}$ is given by

$$L(\boldsymbol{\theta}) = \mathbb{P}_{\boldsymbol{\theta}}(Y_n, \ldots, Y_s|Y_{s-1}, \ldots, Y_0) = \prod_{t=s}^{n} \mathbb{P}_{\boldsymbol{\theta}}(Y_t|Y_{t-s}).$$

The conditional log-likelihood function is then given by

$$\ell(\boldsymbol{\theta}) = \sum_{t=s}^{n} \log[\mathbb{P}_{\boldsymbol{\theta}}(Y_t|Y_{t-s})]. \tag{7}$$

The CML estimator of $\boldsymbol{\theta}$ is the value $\hat{\boldsymbol{\theta}}_{\text{CML}}$ that maximizes $\ell(\boldsymbol{\theta})$. Since $\ell'(\boldsymbol{\theta})$ is a nonlinear function, the maximum likelihood estimate of $\boldsymbol{\theta}$ must be computed using numerical methods. Theorem 4 gives an asymptotic result for the CML estimator of a Poisson INAR(1)$_s$ process.

THEOREM 4 *The CML estimator $\hat{\boldsymbol{\theta}}_{\mathrm{CML}} = (\hat{\phi}_{\mathrm{CML}}, \hat{\lambda}_{\mathrm{CML}})^{\mathsf{T}}$ of the parameter $\boldsymbol{\theta} = (\phi, \lambda)^{\mathsf{T}}$ of a Poisson INAR(1)$_s$ process satisfies*

$$\sqrt{n} \begin{pmatrix} \hat{\phi}_{\mathrm{CML}} - \phi \\ \hat{\lambda}_{\mathrm{CML}} - \lambda \end{pmatrix} \xrightarrow{d} \mathcal{N}(0, \mathbf{I}^{-1}(\boldsymbol{\theta})),$$

where $\mathbf{I}(\boldsymbol{\theta})$ is a 2×2 Fisher information matrix.

The proof of Theorem 2 is omitted here since it is a straightforward consequence for the seasonal case of a result obtained by Franke and Seligmann.[15]

4. Forecasting

Consider the problem of forecasting a future value Y_{n+h}, $h \in \mathbb{N}$, based on the series up to time $n \in \mathbb{N}$. Our starting point is that, for INAR(1)$_s$ process, Equation (1) implies that the distribution of Y_{n+h} can be expressed as

$$Y_{n+h} \stackrel{d}{=} \phi^q \circ Y_{n-r} + \sum_{j=0}^{q-1} \phi^j \circ \epsilon_{n+h-js}, \quad h \in \mathbb{N}, \tag{8}$$

where $q := \lceil h/s \rceil$ and $r := qs - h$, with $\lceil x \rceil$ denoting the upper integer part of $x \in \mathbb{R}$, that is, $\lceil x \rceil := \min\{n \in \mathbb{Z} | x \le n\}$. Then, it is clear that $r \in \{0, \dots, s-1\}$. Thus, from Equation (8), the distribution of Y_{n+h} can be written as the convolution of $\mathcal{B}(Y_{n-r}, \phi^q)$ and $\mathcal{P}o(\lambda(1-\phi^q)/(1-\phi))$ distributions, which is an analogue of a theorem of Freeland [16] for non-seasonal case.

For all $n, h \in \mathbb{N}$, consider to forecast Y_{n+h} by an $\mathcal{F}_n$-measurable random variable with finite second moment $\hat{Y}_n(h)$. Then, the forecast $\hat{Y}_n(h)$ of Y_{n+h} with minimum mean square error $\mathbb{E}_{\boldsymbol{\theta}}(Y_{n+h} - \hat{Y}_n(h))^2$ is achieved by the conditional expectation $\mathbb{E}_{\boldsymbol{\theta}}[Y_{n+h}|\mathcal{F}_n]$. For the INAR(1)$_s$ process, this h-step ahead conditional expectation is given by

$$\mathbb{E}_{\boldsymbol{\theta}}[Y_{n+h}|\mathcal{F}_n] = \phi^q \left(Y_{n-r} - \frac{\lambda}{1-\phi} \right) + \frac{\lambda}{1-\phi}, \quad h \in \mathbb{N}.$$

The following proposition formalizes some properties of the h-step ahead conditional expectation for INAR(1)$_s$ model.

PROPOSITION 5 *Let $\{Y_t\}_{t \in \mathbb{Z}}$ be a stationary Poisson INAR(1)$_s$ process and $n, h \in \mathbb{N}$. Then, the following properties hold*

(1) $Y_{n+h}|\mathcal{F}_n \sim \mathcal{B}(Y_{n-r}, \phi^q) * \mathcal{P}o(\lambda(1-\phi^q)/(1-\phi))$,
(2) $\mathbb{E}_{\boldsymbol{\theta}}[Y_{n+h}|\mathcal{F}_n] = \phi^q(Y_{n-r} - \lambda/(1-\phi)) + \lambda/(1-\phi)$,
(3) $\mathrm{var}_{\boldsymbol{\theta}}[Y_{n+h}|\mathcal{F}_n] = \phi^q(1-\phi^q)Y_{n-r} + \lambda(1-\phi^q)/(1-\phi)$,
(4) $\lim_{h \to \infty} \mathbb{E}_{\boldsymbol{\theta}}[Y_{n+h}|\mathcal{F}_n] = \lim_{h \to \infty} \mathrm{var}_{\boldsymbol{\theta}}[Y_{n+h}|\mathcal{F}_n] = \lambda/(1-\phi)$,

where $q := \lceil h/s \rceil$ and $r := qs - h$.

Proof See the appendix. ∎

Therefore, based on the sample $Y_0, Y_1, \ldots, Y_n$, a forecast $\hat{Y}_n(h), h \in \mathbb{N}$, is obtained by

$$\hat{Y}_n(h) = \hat{\phi}^q \left(Y_{n-r} - \frac{\hat{\lambda}}{1 - \hat{\phi}} \right) + \frac{\hat{\lambda}}{1 - \hat{\phi}}, \quad h \in \mathbb{N}, \tag{9}$$

where $\hat{\lambda}$ and $\hat{\phi}$ are estimators for λ and ϕ, respectively. The forecast error is defined by $\hat{e}_n(h) := Y_{n+h} - \hat{Y}_n(h)$. Similar to the INAR(1) models, the forecast computed by Equation (9) will seldom produce integer-valued $\hat{Y}_n(h)$. To circumvent this problem, Freeland and McCabe [17] suggested to use, instead of the conditional expectation, the conditional median, which minimizes the expected absolute error. However, the conditional median, for most situations, cannot be easily computed.

5. Monte Carlo simulation study

The asymptotic properties of the estimators discussed in the previous sections are now investigated for finite sample sizes $n = 100, 200, 500$ from Poisson INAR(1)$_s$ series with $s = 12$ and $\phi = 0.3, 0.5, 0.8$, $\{\epsilon_t\}_{t \in \mathbb{Z}}$ being an i.i.d. Poisson sequence with mean $\lambda = 1, 5$. The samples were simulated using the R programming language.[18] For comparison purposes, the Yule–Walker (YW) approach was also considered in the empirical investigation, since it is a widely used estimation method in time-series models. The YW estimators of ϕ and λ are based upon the sample ACF $\hat{\rho}$, using that $\rho(s) = \phi$, and the first moment of Y_t, which is $\mathbb{E}(Y_t) = \lambda/(1 - \phi)$. They are given by

$$\hat{\phi}_{\text{YW}} = \hat{\rho}(s) := \frac{\sum_{t=1}^{n-s}(Y_t - \bar{Y})(Y_{t+s} - \bar{Y})}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2}, \quad \hat{\lambda}_{\text{YW}} := (1 - \hat{\phi}_{\text{YW}})\bar{Y}, \tag{10}$$

where $\bar{Y} := (1/n)\sum_{t=1}^{n} Y_t$ denotes the sample mean. It is well-known that the estimators above are strongly consistent.[19]

The empirical results displayed in the tables, that is, the empirical biases and mean square errors (MSE), were computed over 1000 replications. The values of the MSE are given between parenthesis. The CML estimates of ϕ and λ were obtained using the BFGS (Broyden, Fletcher, Goldfarb, Shanno) quasi-Newton nonlinear optimization algorithm with numerical derivatives.

From Tables 1 and 2, it can be seen that the CML estimator presents much smaller biases (in absolute values) and MSEs than the other estimators, for all models. As expected, increasing the sample size reduces substantially both bias and MSE. A comparison between Tables 1 and 2 indicates that both bias and MSE are larger in the latter case, which is not a surprising evidence since in the second case, the model has larger mean and variance than in the first one. Another result is related to the size of ϕ. In general, for YW and CLS methods, increasing ϕ, the bias and MSE also increase. This indicates that these two estimation methods are sensitive to a process that is closer to the non-stationary boundary, that is, the model is more near a unit root seasonal INAR process.

For small sample sizes, in general, both bias and MSE for the YW estimators are smaller than those for the CLS method. This may be explained by the fact that the YW estimator is calculated using a sample size of n (Equation (10)), while the CLS estimator is based on a sample size of $n - s$ (Equation (6)).

The empirical investigation presented here suggests that the performance of the CML estimator is much superior to those of the YW and CLS estimators. The superiority of CML was expected, since this estimator uses the whole information of the distribution. However, what the empirical results show is that there is a large degree of superiority of the CML with respect to the other methods.

Table 1.　Biases of estimators for $\lambda = 1$ (MSE in parentheses).

n	ϕ	Bias($\hat{\phi}$)/MSE($\hat{\phi}$)			Bias($\hat{\lambda}$)/MSE($\hat{\lambda}$)		
		YW	CLS	CML	YW	CLS	CML
	0.30	−0.0178	−0.0307	−0.0067	0.0315	0.0508	0.0040
		(0.0125)	(0.0133)	(0.0114)	(0.0353)	(0.0365)	(0.0291)
100	0.50	−0.0240	−0.0334	−0.0081	0.0500	0.0691	0.0064
		(0.0100)	(0.0116)	(0.0063)	(0.0549)	(0.0560)	(0.0304)
	0.80	−0.0267	−0.0362	−0.0031	0.1385	0.1854	0.0113
		(0.0058)	(0.0078)	(0.0012)	(0.1583)	(0.1921)	(0.0289)
	0.30	−0.0115	−0.0156	−0.0067	0.0221	0.0282	0.0115
		(0.0045)	(0.0044)	(0.0035)	(0.0130)	(0.0133)	(0.0104)
250	0.50	−0.0106	−0.0146	−0.0029	0.0254	0.0337	0.0057
		(0.0037)	(0.0040)	(0.0023)	(0.0174)	(0.0184)	(0.0109)
	0.80	−0.0143	−0.0166	−0.0016	0.0700	0.0823	0.0028
		(0.0019)	(0.0022)	(0.0004)	(0.0511)	(0.0572)	(0.0113)
	0.30	−0.0058	−0.0079	−0.0023	0.0063	0.0093	−0.0008
		(0.0022)	(0.0022)	(0.0018)	(0.0055)	(0.0056)	(0.0045)
500	0.50	−0.0033	−0.0056	−0.0007	0.0102	0.0148	0.0029
		(0.0018)	(0.0018)	(0.0010)	(0.0087)	(0.0089)	(0.0052)
	0.80	−0.0086	−0.0098	−0.0003	0.0468	0.0500	0.0043
		(0.0009)	(0.0009)	(0.0002)	(0.0237)	(0.0255)	(0.0055)

Table 2.　Biases of estimators in the case $\lambda = 5$ (MSE in parentheses).

n	ϕ	Bias($\hat{\phi}$)/MSE($\hat{\phi}$)			Bias($\hat{\lambda}$)/MSE($\hat{\lambda}$)		
		YW	CLS	CML	YW	CLS	CML
	0.30	−0.0180	−0.0855	−0.0083	0.2048	0.6947	0.0765
		(0.0113)	(0.0216)	(0.0103)	(0.6722)	(1.3055)	(0.5793)
100	0.50	−0.0199	−0.0933	−0.0033	0.2620	1.0068	0.0335
		(0.0097)	(0.0236)	(0.0060)	(1.1078)	(2.6607)	(0.6218)
	0.80	−0.0304	−0.1509	−0.0003	0.7953	3.7812	0.0017
		(0.0055)	(0.0516)	(0.0010)	(3.5483)	(32.314)	(0.6105)
	0.30	−0.0082	−0.0300	−0.0034	0.0750	0.2348	0.0197
		(0.0041)	(0.0054)	(0.0035)	(0.02340)	(0.03073)	(0.1940)
250	0.50	−0.0083	−0.0307	−0.0025	0.1063	0.3314	0.0262
		(0.0035)	(0.0049)	(0.0021)	(0.03669)	(0.5212)	(0.2121)
	0.80	−0.0135	−0.0406	−0.0011	0.3572	1.7114	0.0305
		(0.0019)	(0.0044)	(0.0003)	(1.2000)	(2.7742)	(0.2269)
	0.30	−0.0019	−0.0125	−0.0002	0.0304	0.1068	0.0077
		(0.0019)	(0.0022)	(0.0017)	(0.1083)	(0.1236)	(0.0905)
500	0.50	−0.0029	−0.0134	0.0009	0.0369	0.1426	−0.0109
		(0.0016)	(0.0019)	(0.0001)	(0.1681)	(0.1982)	(0.1059)
	0.80	−0.0050	−0.0162	0.0000	0.1353	0.4147	0.0001
		(0.0009)	(0.0013)	(0.0002)	(0.5409)	(0.7884)	(0.1104)

6.　Real data example

The model, estimation and forecast methods proposed in this paper are now used to model and forecast the monthly counts of claims of short-term disability benefits. In the data set, all the claimants are male, between the ages of 35 and 54 years, work in the logging industry and reported their claim to the Richmond, BC Workers Compensation Board. Only claimants whose injuries were due to cuts and lacerations were included in the data set. The data consist of 120 observations starting from January 1985 and ending in December 1994. This series was previously studied by Freeland [16] and Zhu and Joe.[12] In the former study, covariates were used
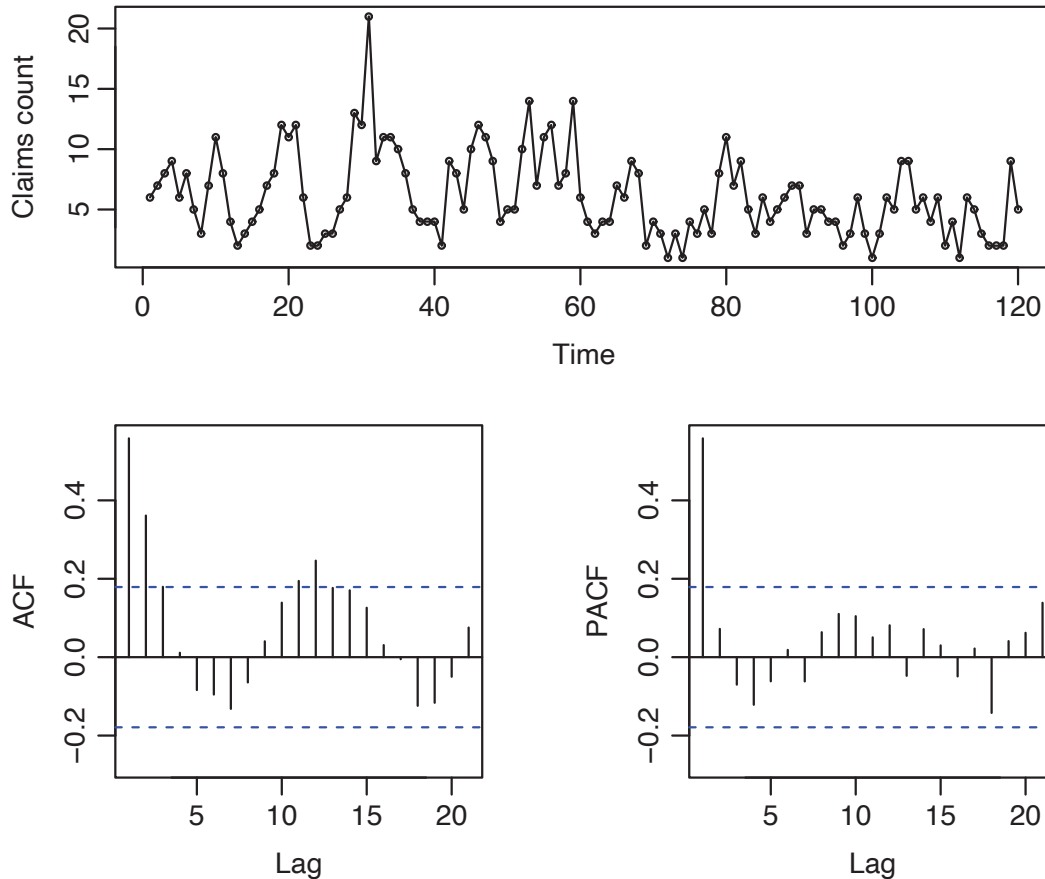
Figure 2. The time-series plot, ACF and PACF of the claims series from 1985 to 1994.

to account for the seasonality feature. In the same direction, the latter also considered covariates functions to explain seasonality in counting models with Poisson and negative binomial marginals to adjust the series. These models are not stationary. Zhu and Joe [12] also fitted the series using INAR(1) processes with Poisson and negative binomial innovations without consideration of seasonality. As was expected, the models with covariates, that is, the non-stationary processes, produced smaller Akaike information criterion (AIC) values, however with larger parametric dimensions.

Since the INAR(1)$_s$ process is a simple stationary model and accounts for seasonal patterns, the above applications motivated its use as a candidate model to fit the monthly counts of injured workers.

The series together with its sample ACF and partial autocorrelation function (PACF) are displayed in Figure 2. The plot of the series indicates that it is a mean stationary time series with an apparent seasonal and the serial correlations behaviours. These phenomena can also be observed in the plot of the sample ACF and the PACF. The plot shows the geometric decrease in the ACF with a seasonal period of 12 and, also, the geometric decrease in the serial correlations. The seasonal feature may be due to the fact that the logging industry is more active in the warmer months. The behaviour of PACF may justify also the use of counting time-series models without consideration of seasonality. Therefore, for comparison purposes, these data were also fitted by an INAR(1) model. The first 110 observations were used to model the series, while the remaining 10 observations were considered for forecasting purposes.

Table 3 provides the CML and CLS estimates (with standard errors in parentheses) of the model parameters and two goodness of-fit statistics: the AIC and the Bayesian information criterion (BIC).

Table 3. Estimated parameters (standard errors in parentheses), AIC and BIC.

Model		CML estimates	CLS estimates	AIC	BIC
INAR(1)$_{12}$	$\hat{\phi}$	0.1746(0.0036)	0.2410(0.0899)	530.613	536.013
	$\hat{\lambda}$	5.1391(0.1951)	4.7554(0.5897)		
INAR(1)	$\hat{\phi}$	0.4418(0.0029)	0.5510(0.0783)	538.469	543.869
	$\hat{\lambda}$	3.5224(0.1364)	2.8526(0.5079)		
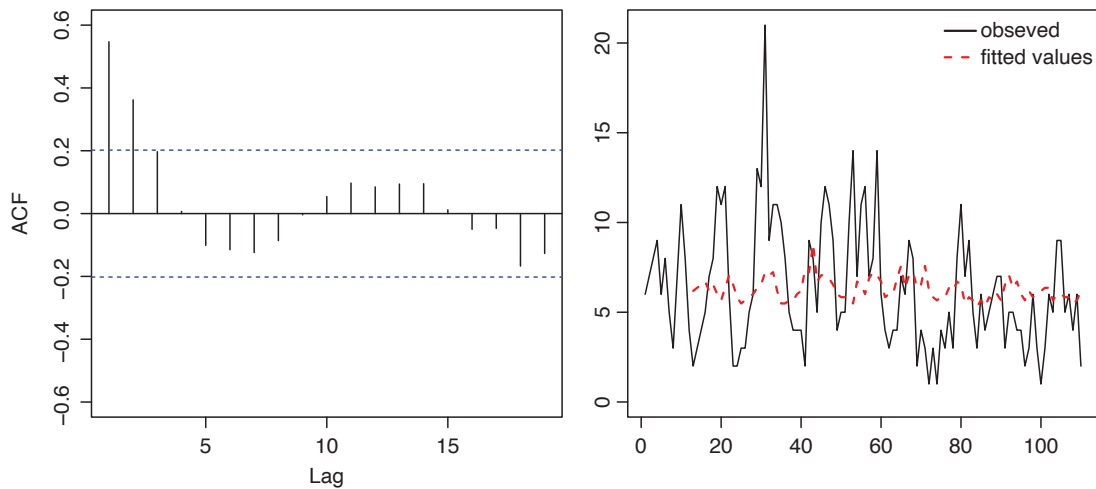
Figure 3. Sample autocorrelations of the residuals obtained from INAR(1)$_{12}$ model and fitted values.

The results displayed in Table 3 suggest that the INAR(1)$_{12}$ model produces better fits to the data than the INAR(1) model. The model fitted by CML estimation is

$$Y_t = 0.1746 \circ Y_{t-12} + \epsilon_t$$

with $\epsilon_t \sim \mathcal{P}o(5.1391)$.

The sample autocorrelations of the residuals from the INAR(1)$_{12}$ model and the fitted values are shown in Figure 3. Although the estimated results in Table 3 indicated that the INAR(1)$_{12}$ model was more adequate to fit the data, there is still serial correlation in the residuals. Therefore, the proposed model was not able to capture the total correlation in the data. This empirical evidence suggests a more complete model which can capture both the seasonal and non-seasonal correlations. One suggestion may be to use an INAR model that incorporates both seasonal and serial correlations. This issue is beyond this paper.

The forecast investigation is given in Table 4. The point predictions of seasonal and non-seasonal models are close to each other and they both seem to be reasonable estimates of the future h-step ahead observations. The interval prediction (IP) is obtained using the result of Silva et al.,[20, Section 3] for the seasonal and non-seasonal cases. For more details, see Silva et al.[20] It can be seen that, in most cases, intervals obtained by the INAR(1)$_{12}$ model have smaller widths. The root mean square forecast errors (RMSE), at the bottom of the Table 4, shows average fitting in both cases are very similar, being the non-seasonal the one with smaller value. It should be remarked that the fitting is a little bit better in the non-seasonal case, contrary to the goodness-of-fit based on different information criteria in Table 3. This forecasting investigation may also corroborate to the use of a more complete model, that is, a model with seasonal and serial correlations, to fit the data.

Table 4. $100(1 - \gamma)\%$ Prediction intervals with actual values and forecasting values for the burns data.

h	Y_{110+h}	Point prediction		IP ($\gamma = 0.05$)		IP ($\gamma = 0.10$)	
		$INAR(1)_{12}$	$INAR(1)$	$INAR(1)_{12}$	$INAR(1)$	$INAR(1)_{12}$	$INAR(1)$
1	4	5.663	4.406	(0,10)	(0,8)	(0,9)	(0,7)
2	1	5.314	5.469	(0,9)	(0,10)	(0,8)	(0,9)
3	6	5.663	5.939	(0,10)	(0,11)	(0,9)	(0,9)
4	5	6.187	6.146	(0,10)	(0,10)	(0,9)	(0,9)
5	3	6.012	6.238	(0,10)	(0,11)	(0,9)	(0,10)
6	2	6.711	6.278	(0,11)	(0,11)	(0,10)	(0,10)
7	2	6.711	6.296	(0,11)	(0,11)	(0,10)	(0,10)
8	2	6.012	6.304	(0,10)	(0,11)	(0,9)	(0,10)
9	9	6.187	6.308	(0,10)	(0,11)	(0,9)	(0,10)
10	5	5.838	6.309	(0,10)	(0,11)	(0,9)	(0,10)
RMSE		3.179	3.101				

7. Conclusions

In this paper, the non-negative INAR process with seasonal structure based on the model proposed by Al-Osh and Alzaid [3] is introduced and its main properties are derived. Three estimators for the model parameters are considered, the YW, conditional least squares and CML estimators, and their relative merits as estimators are compared from simulation studies. The CML estimator, as expected, is the most efficient, although it is also the most computationally intensive of the three. However, the magnitude of the gain in terms of bias and MSE makes the CML estimator a much better choice for seasonal models in all situations, a similar conclusion to that obtained for the standard INAR(1) model; see, for example, Al-Osh and Alzaid.[3] Thus, the use of CML to estimate the model parameters of an $INAR(1)_s$ process is recommended here. Finally, the $INAR(1)_s$ model is applied to a real data set to show the potential of this new model. The INAR(1) model is also used to fit the same data. Not surprisingly, the $INAR(1)_s$ model presents a better fit. Looking at the ACF plot, it is clear that there are seasonal and nonseasonal autoregressive contributions. Therefore, an interesting extension of the model proposed here is to combine both INAR(1) and $INAR(1)_s$ models in only one process, such as the standard SARMA (Seasonal Autorregressive Moving Average) time-series models. This is a current research of the authors and the subject of a forthcoming paper.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

[1] Box GEP, Jenkins GM, Reinsel GC. Time series analysis, forecasting and control. 4th ed. Hoboken, NJ: Wiley; 2008.

[2] Mckenzie E. Some ARMA models for dependent sequences of Poisson counts. Adv Appl Probab. 1988;20:822–835.

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

[4] Weiß CH. Thinning operations for modelling time series of counts–a survey. Adv Stat Anal. 2008;92:319–341.

[5] Jung RC, Ronning G, Tremayne AR. Estimation in conditional first order autoregression with discrete support. Statist Papers. 2005;46:195–224.

[6] Freeland RK, McCabe BPM. Asymptotic properties of CLS estimators in the Poisson AR(1) model. Statist Probab Lett. 2005;73:147–153.

[7] Monteiro M, Scotto MG, Pereira I. Integer-valued autoregressive processes with periodic structure. J Statist Plann Inference. 2010;140:1529–1541.

[8] Moriña D, Puig P, Ríos J, Vilella A, Trilla A. A statistical model for hospital admissions caused by seasonal diseases. Stat Med. 2011;30:3125–3136.

[9] Freeland RK, McCabe BPM. Analysis of low count time series data by Poisson autoregression. J Time Ser Anal. 2004;25:701–722.

[10] Quddus MA. Time series count data models: an empirical application to traffic accidents. Accid Anal Preven. 2008;40:1732–1741.

[11] Zhang H, Wang D, Zhu F. Inference for INAR(p) processes with signed generalized power series thinning operator. J Stat Plann Inference. 2010;140:667–683.

[12] Zhu R, Joe H. Modelling count data time series with Markov processes based on binomial thinning. J Time Ser Anal. 2006;27:725–738.

[13] Steutel FW, van Harn K. Discrete analogues of self-decomposability and stability. Ann Probab. 1979;7:893–899.

[14] Klimko LA, Nelson PI. On conditional least squares estimation for stochastic processes. Ann Statist. 1978;6:629–642.

[15] Franke J, Seligmann T. Conditional maximum likelihood estimates for INAR(1) processes and their application to modelling epileptic seizure counts. In: Subba Rao T, editor. Developments in time series analysis. London: Chapman-Hall; 1993. p. 310–330.

[16] Freeland RK. Statistical analysis of discrete-time series with applications to the analysis of workers compensation claims data [Ph.D. thesis]. Canada: University of British Columbia; 1998.

[17] Freeland RK, McCabe BPM. Forecasting discrete valued low count time series. Int J Forecast. 2004;20:427–434.

[18] R Development Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2009. Available from http://www.R-project.org.

[19] Du J, Li Y. The integer-valued autoregressive (INAR(p)) model. J Time Ser Anal. 1991;12:129–142.

[20] Silva N, Pereira I, Silva ME. Forecasting in INAR(1) model. REVSTAT. 2009;7:119–134.

[21] Silva ME, Oliveria VL. Difference equations for the higher-order moments and cumulants of the INAR(1) model. J Time Ser Anal. 2004;25:317–333.

Appendix

Proof of Proposition 1

Proof Define the random variables $X_{t,n}$ as

$$X_{t,n} := \sum_{k=0}^{n} \phi^k \circ \epsilon_{t-ks} = \epsilon_t + \sum_{k=1}^{n} \sum_{j=1}^{\epsilon_{t-ks}} Z_{t,k,j}, \quad t \in \mathbb{Z}, \ n \in \mathbb{Z}_+, \tag{A1}$$

where the Bernoulli variables $\{Z_{t,k,j}\}_{k,j \in \mathbb{N}}$ are independent and mutually independent of the innovation process $\{\epsilon_t\}_{t \in \mathbb{Z}}$ with $\mathbb{E}Z_{t,k,j} = \phi^k$ for all $k, j \in \mathbb{N}$ and $t \in \mathbb{Z}$. We prove that $\{X_{t,n}\}_{n \in \mathbb{Z}_+}$ forms a Cauchy sequence in probability for all $t \in \mathbb{Z}$. If λ and σ^2 represent, respectively, the mean and variance of the innovation sequence, then we obtain, for all $0 < n < m$,

$$\mathbb{E}\left[\sum_{k=n}^{m} \phi^k \circ \epsilon_{t-ks} \right] = \sum_{k=n}^{m} \phi^k \lambda \leq \frac{\phi^n}{1-\phi} \lambda,$$

$$\mathrm{var}\left[\sum_{k=n}^{m} \phi^k \circ \epsilon_{t-ks} \right] = \sum_{k=n}^{m} (\phi^{2k}\sigma^2 + \phi^k(1-\phi^k)\lambda) \leq \frac{\phi^{2n}}{1-\phi^2}\sigma^2 + \frac{\phi^n}{1-\phi}\lambda, \tag{A2}$$

using properties (iv) and (v) of the thinning operator, see Silva and Oliveria.[21] Since $\phi \in [0,1)$, the right-hand sides in Equation (12) tend to 0 as $n \to \infty$. This implies that the sequence $\{X_{t,n}\}_{n \in \mathbb{Z}_+}$ forms a Cauchy sequence in mean square sense and hence in probability, as well. Thus, for all $t \in \mathbb{Z}$, there exists a random variable X_t, which is the limit on the right-hand side in Equation (3), such that $X_{t,n} \overset{P}{\to} X_t$ as $n \to \infty$. Let the non-negative integer-valued stochastic process $\{Y_t\}_{t \in \mathbb{Z}}$ satisfy the Equation (1). By substituting successively we obtain

$$Y_t = \phi \circ Y_{t-s} + \epsilon_t = \phi \circ (\phi \circ Y_{t-2s} + \epsilon_{t-s}) + \epsilon_t.$$

Using that $\phi \circ (\psi \circ X) \overset{d}{=} (\phi \psi) \circ X$ and $\phi \circ (X + Y) \overset{d}{=} \phi \circ X + \phi \circ Y$ for any $\phi, \psi \in [0,1]$ and any independent pair of non-negative integer-valued random variables X, Y (see [3, p.262]), we have the following equality in distribution:

$$Y_t \overset{d}{=} \phi^2 \circ Y_{t-2s} + \phi \circ \epsilon_{t-s} + \epsilon_t = \phi^2 \circ Y_{t-2s} + X_{t,1},$$

since ϵ_{t-s} and Y_{t-2s} are independent. By induction, for all $n \in \mathbb{N}$, we have

$$Y_t \overset{d}{=} \phi^n \circ Y_{t-ns} + \sum_{k=0}^{n-1} \phi^k \circ \epsilon_{t-ks} = \phi^n \circ Y_{t-ns} + X_{t,n-1}. \tag{A3}$$

If μ_Y and σ_Y^2 represent, respectively, the mean and variance of a stationary solution $\{Y_t\}_{t \in \mathbb{Z}}$, we obtain

$$\mathbb{E}[\phi^n \circ Y_{t-ns}] = \phi^n \mu_Y,$$

$$\mathrm{var}[\phi^n \circ Y_{t-ns}] = \phi^{2n} \sigma_Y^2 + \phi^n (1 - \phi^n) \mu_Y.$$

Since $\phi \in [0,1)$, we obtain $\lim_{n \to \infty} \mathbb{E}[\phi^n \circ Y_{t-ns}] = \lim_{n \to \infty} \mathrm{var}[\phi^n \circ Y_{t-ns}] = 0$. Therefore, $(\phi^n \circ Y_{t-ns}) \overset{P}{\to} 0$ as $n \to \infty$, and thus, by Equation (13), $X_{t,n} \overset{d}{\to} Y_t$ as $n \to \infty$ for all $t \in \mathbb{Z}$, where $\overset{d}{\to}$ denotes convergence in distribution. Hence, $Y_t \overset{d}{=} X_t$ for all $t \in \mathbb{Z}$, which means the uniqueness of the stationary marginal solution. Finally, it is showed that the distribution of the process $\{X_t\}_{t \in \mathbb{Z}}$ is the solution of the Equation (1). Using the above-mentioned properties of the binomial thinning operator, the following is derived:

$$X_{t,n} = \sum_{k=0}^{n} \phi^k \circ \epsilon_{t-ks} \overset{d}{=} \epsilon_t + \phi \circ \left(\sum_{k=0}^{n-1} \phi^k \circ \epsilon_{t-s-ks} \right) = \phi \circ X_{t,n-1} + \epsilon_t, \quad t \in \mathbb{Z}.$$

Taking the limit in probability as $n \to \infty$ the equality $X_t \overset{d}{=} \phi \circ X_t + \epsilon_t$ is obtained, that is, the distribution of X_t is a stationary marginal distribution to Equation (1). This completes the proof. ∎

Proof of Proposition 2

Proof (a) Let φ_W denote the characteristic function of a random variable W. If $W = \phi \circ X$, where X is a non-negative integer-valued random variable, then $\varphi_W(t) = \mathbb{E}(\phi \, e^{it} + 1 - \phi)^X$. In particular, if $X \sim Po(\lambda)$, then $\varphi_W(u) = \exp\{\phi \lambda (e^{iu} - 1)\}$. By the uniqueness of the characteristic function, using Equation (13), $\varphi_{Y_t} = \varphi_{\phi^n \circ Y_{t-ns}} \cdot \varphi_{X_{t,n-1}}$ is obtained. Since $(\phi^n \circ Y_{t-ns}) \overset{P}{\to} 0$ as $n \to \infty$, (see the proof of Proposition 1) we have $\varphi_{\phi^n \circ Y_{t-ns}} \to 1$ as $n \to \infty$. Thus, $\varphi_{X_{t,n}} \to \varphi_{Y_t}$ as $n \to \infty$. Since ϵ_t has Poisson distribution with mean λ for all $t \in \mathbb{Z}$, $X_{t,n}$ is a sum of independent random variables and $\phi \in [0,1)$, we obtain

$$\varphi_{X_{t,n}}(u) = \prod_{k=0}^{n} \exp\{\phi^k \lambda (e^{iu} - 1)\} = \exp \left\{ \sum_{k=0}^{n} \phi^k \lambda (e^{iu} - 1) \right\} \to \exp \left\{ \frac{\lambda}{1 - \phi} (e^{iu} - 1) \right\}$$

as $n \to \infty$ for all $u \in \mathbb{R}$. Thus, by the uniqueness of the characteristic function, $Po(\lambda/(1 - \phi))$ provides the only stationary distribution.

(b) Let γ denote the covariance function of the stationary INAR(1)$_s$ process $\{Y_t\}_{t \in \mathbb{Z}}$, that is, $\gamma(k) := \mathrm{cov}(Y_{t-k}, Y_t)$, $k \in \mathbb{Z}$. By Equation (1), we have

$$\gamma(k) = \mathrm{cov}(Y_{t-k}, Y_t) = \mathrm{cov}(Y_{t-k}, \phi \circ Y_{t-s}) + \mathrm{cov}(Y_{t-k}, \epsilon_t) = \phi \, \mathrm{cov}(Y_{t-k}, Y_{t-s}) = \phi \gamma(k - s),$$

where we used the assumption that Y_t is independent of future values of ϵ_t and the fact that $\mathrm{cov}(\phi \circ U, V) = \phi \, \mathrm{cov}(U, V)$ if the Bernoulli variables defining $\phi \circ U$ are mutually independent both of U and V. Suppose that $k = hs + r$, where $h \in \mathbb{Z}_+$ and $r \in \{0, 1, \ldots s - 1\}$. Then, by induction, we obtain $\gamma(k) = \phi^h \gamma(r)$. Thus, if $r = 0$, that is, k is a positive integer multiple of s, then $\gamma(k) = \phi^h \gamma(0)$, where $\gamma(0)$ is the variance of the process $\{Y_t\}_{t \in \mathbb{Z}}$. If $r > 0$, then it is clear that it is enough to compute the covariance between Y_t and Y_{t-r}. However, these random variables belong to the different seasonal components of the INAR(1)$_s$ process and thus they are independent. This gives the result. ∎

Proof of Theorem 3

Proof   Let $Y_0, Y_1, \ldots, Y_n$ be a sample of a Poisson INAR(1)$_s$ process. The aim of the following discussion is to show that the regularity conditions given in Theorem 3.2 of [14] are satisfied. Since $\partial g/\partial \phi = y$, $\partial g/\partial \lambda = 1$, and $\partial^2 g/\partial \phi^2 = \partial^2 g/\partial \phi \partial \lambda = \partial^2 g/\partial \lambda^2 = 0$, the regularity conditions (i), (ii), and (iii) on Klimko and Nelson [14, p.634] hold. Define the 2×2 matrix V according to Equation (3.2) in [14] as

$$V_{ij} = \mathbb{E}_{\boldsymbol{\theta}} \left(\frac{\partial g(\boldsymbol{\theta}, Y_{t-s})}{\partial \theta_i} \cdot \frac{\partial g(\boldsymbol{\theta}, Y_{t-s})}{\partial \theta_j} \right), \quad i, j = 1, 2$$

and the 2×2 matrix W according to Equation (3.5) in [14] as

$$W_{ij} = \mathbb{E}_{\boldsymbol{\theta}} \left(u_t^2(\boldsymbol{\theta}) \frac{\partial g(\boldsymbol{\theta}, Y_{t-s})}{\partial \theta_i} \cdot \frac{\partial g(\boldsymbol{\theta}, Y_{t-s})}{\partial \theta_j} \right), \quad i, j = 1, 2,$$

where $u_t(\boldsymbol{\theta}) = Y_t - g(\boldsymbol{\theta}, Y_{t-s})$. Using the above-mentioned partial derivatives of the function $g(\cdot, \cdot)$,

$$V_{11} = \mathbb{E}_{\boldsymbol{\theta}}(Y_{t-s}^2) = \frac{\lambda}{1-\phi} + \frac{\lambda^2}{(1-\phi)^2},$$

$$V_{12} = \mathbb{E}_{\boldsymbol{\theta}}(Y_{t-s}) = \frac{\lambda}{1-\phi},$$

$$V_{22} = \mathbb{E}_{\boldsymbol{\theta}}(1) = 1$$

and

$$W_{11} = \mathbb{E}_{\boldsymbol{\theta}}((Y_t - \phi Y_{t-s} - \lambda)^2 Y_{t-s}^2) = \phi\lambda + \lambda^2 \left(\frac{1+3\phi}{1-\phi} \right) + \lambda^3 \left(\frac{1+\phi}{(1-\phi)^2} \right),$$

$$W_{12} = \mathbb{E}_{\boldsymbol{\theta}}((Y_t - \phi Y_{t-s} - \lambda)^2 Y_{t-s}) = \phi\lambda + \frac{1+\phi}{1-\phi}\lambda^2,$$

$$W_{22} = \mathbb{E}_{\boldsymbol{\theta}}((Y_t - \phi Y_{t-s} - \lambda)^2) = \lambda(1+\phi)$$

are obtained. The entries of the matrix W are derived by the following argument. Since, by Equation (1),

$$Y_t - \phi Y_{t-s} - \lambda = \sum_{j=1}^{Y_{t-s}} (Z_j - \phi) + (\epsilon_t - \lambda),$$

where $\{Z_j\}_{j \in \mathbb{N}}$ is an i.i.d. Bernoulli random sequence with mean ϕ, then, by conditioning with respect to Y_{t-s}, we obtain

$$W_{11} = \mathbb{E}_{\boldsymbol{\theta}} \left(\left(\sum_{j=1}^{Y_{t-s}} (Z_j - \phi) + (\epsilon_t - \lambda) \right)^2 Y_{t-s}^2 \right)$$

$$= \mathbb{E}_{\boldsymbol{\theta}} \left(Y_{t-s}^2 \sum_{j=1}^{Y_{t-s}} (Z_j - \phi)^2 \right) + \mathbb{E}_{\boldsymbol{\theta}} \left((\epsilon_t - \lambda)^2 \right) \mathbb{E}_{\boldsymbol{\theta}}(Y_{t-s}^2).$$

Here, again by conditioning with respect to Y_{t-s}, we have

$$\mathbb{E}_{\boldsymbol{\theta}} \left(Y_{t-s}^2 \sum_{j=1}^{Y_{t-s}} (Z_j - \phi)^2 \right) = \mathbb{E}_{\boldsymbol{\theta}}(Y_{t-s}^3) \mathbb{E}_{\boldsymbol{\theta}}((Z_j - \phi)^2)$$

and the result follows for entry W_{11} since $\mathbb{E}_{\boldsymbol{\theta}}(Y_{t-s}^3) = \lambda/(1-\phi) + 3(\lambda/(1-\phi))^2 + (\lambda/(1-\phi))^3$. The entries W_{12} and W_{22} can be similarly derived. It is easy to check that the matrix V is positive definite. Finally, it is clear that conditions

(3.3) and (3.4) in [14] are also satisfied. Hence, the estimator $\hat{\boldsymbol{\theta}}_{\text{CLS}}$ of $\boldsymbol{\theta}$ has the following asymptotic distribution:

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_{\text{CLS}} - \boldsymbol{\theta}) \xrightarrow{d} \mathcal{N}(0, V^{-1}WV^{-1}).$$

Since

$$V^{-1} = \begin{bmatrix} \dfrac{1-\phi}{\lambda} & -1 \\ -1 & 1 + \dfrac{\lambda}{1-\phi} \end{bmatrix}, \quad W = \begin{bmatrix} \phi\lambda + \lambda^2\left(\dfrac{1+3\phi}{1-\phi}\right) + \lambda^3\left(\dfrac{1+\phi}{(1-\phi)^2}\right) & \phi\lambda + \dfrac{1+\phi}{1-\phi}\lambda^2 \\ \phi\lambda + \dfrac{1+\phi}{1-\phi}\lambda^2 & (1+\phi)\lambda \end{bmatrix},$$

$$V^{-1}WV^{-1} = \begin{bmatrix} \dfrac{\phi(1-\phi)^2}{\lambda} + (1-\phi^2) & -(1+\phi)\lambda \\ -(1+\phi)\lambda & \lambda + \dfrac{1+\phi}{1-\phi}\lambda^2 \end{bmatrix}$$

is obtained and the proof is finished. ∎

Proof of Proposition 5

Proof (1) Let $q := \lceil h/s \rceil$ and $r := qs - h$. Then, from Equation (A3), Equation (8) is obtained, that is, $Y_{t+h} = Y_{t-r+qs} = \phi^q \circ Y_{t-r} + X_{t-r,q-1}$, see Equation (A1). This statement follows from the facts that $\phi^q \circ Y_{t-r} \sim \mathcal{B}(Y_{n-r}, \phi^q)$, $X_{t-r,q-1} \sim \mathcal{P}o(\lambda(1-\phi^q)/(1-\phi))$ and the independence of $\phi^q \circ Y_{t-r}$ and $X_{t-r,q-1}$.

(2) Using Equation (8), $\mathbb{E}(\phi \circ X|X) = \phi X$, where X is a non-negative integer-valued random variable, and the fact that ϵ_{n+h-js} is independent of the σ-algebra $\mathcal{F}_n$ for all $j = 0, 1, \ldots, q-1$, $\mathbb{E}_{\boldsymbol{\theta}}(Y_{n+h}|\mathcal{F}_n)$ is computed as follows:

$$\mathbb{E}_{\boldsymbol{\theta}}(Y_{n+h}|\mathcal{F}_n) = \mathbb{E}_{\boldsymbol{\theta}}\left(\phi^q \circ Y_{n-r} + \sum_{j=0}^{q-1}\phi^j \circ \epsilon_{n+h-js} \,\middle|\, \mathcal{F}_n\right)$$

$$= \phi^q Y_{n-r} + \lambda\sum_{j=0}^{q-1}\phi^j$$

$$= \phi^q\left(Y_{n-r} - \frac{\lambda}{1-\phi}\right) + \frac{\lambda}{1-\phi}.$$

(3) Using Equation (8), $\text{var}(\phi \circ X|X) = \phi(1-\phi)X$, where X is a non-negative integer-valued random variable, and the fact that ϵ_{n+h-js} is independent of the σ-algebra $\mathcal{F}_n$ for all $j = 0, 1, \ldots, q-1$, $\text{var}_{\boldsymbol{\theta}}(Y_{n+h}|\mathcal{F}_n)$ is computed as follows:

$$\text{var}_{\boldsymbol{\theta}}(Y_{n+h}|\mathcal{F}_n) = \text{var}_{\boldsymbol{\theta}}\left(\phi^q \circ Y_{n-r} + \sum_{j=0}^{q-1}\phi^j \circ \epsilon_{n+h-js} \,\middle|\, \mathcal{F}_n\right)$$

$$= \text{var}_{\boldsymbol{\theta}}(\phi^q \circ Y_{n-r}|Y_{n-r}) + \text{var}_{\boldsymbol{\theta}}\left(\sum_{j=0}^{q-1}\phi^j \circ \epsilon_{n+h-js}\right)$$

$$= \phi^q(1-\phi^q)Y_{n-r} + \sum_{j=0}^{q-1}\text{var}_{\boldsymbol{\theta}}(\phi^j \circ \epsilon_{n+h-js})$$

$$= \phi^q(1-\phi^q)Y_{n-r} + \frac{\lambda(1-\phi^q)}{1-\phi}.$$

(4) Using the results given in (2) and (3), the following limits are obtained:

$$\lim_{h\to\infty}\mathbb{E}_{\boldsymbol{\theta}}(Y_{n+h}|\mathcal{F}_n) = \lim_{h\to\infty}\left[\phi^q\left(Y_{n-r} - \frac{\lambda}{1-\phi}\right) + \frac{\lambda}{1-\phi}\right] = \frac{\lambda}{1-\phi},$$

$$\lim_{h\to\infty}\text{var}_{\boldsymbol{\theta}}(Y_{n+h}|\mathcal{F}_n) = \lim_{h\to\infty}\left[\phi^q(1-\phi^q)Y_{n-r} + \frac{\lambda(1-\phi^q)}{1-\phi}\right] = \frac{\lambda}{1-\phi}.$$

∎