Shahid Chamran University of Ahvaz

Journal of Advanced Mathematical Modeling
ISSN (print): 2251-8088, ISSN (on-line): 2645-6141
Vol. 14, No. 3 (2024), pp. 16-40.
DOI: 10.22055/jamm.2024.45842.2246

A SEASONAL INTEGER-VALUED AR(1) MODEL WITH DELAPORTE MARGINAL DISTRIBUTION

MARYAM SHALBAF*, GHOLAMALI PARHAM AND RAHIM CHINIPARDAZ

Communicated by: Gholamreza Mohtashami

Abstract. Real-count data time series often show the phenomenon of overdispersion. This paper introduces the first-order integer-valued autoregressive process with a seasonal structure. The univariate marginal distribution is derived from the Delaporte distribution. The innovations are the convolution of Poisson with α-fold zero modified geometric distribution, based on the binomial thinning operator, for modeling integer-valued time series with overdispersion. Some properties of the model are derived. The methods of Yule-Walker, conditional least squares, and conditional maximum likelihood are used for estimating the parameters. The Monte Carlo experiment is conducted to evaluate the performances of these estimators in finite samples. At the end this model is illustrated using a real data set and is compared to some INAR(1) models.

1. Introduction

Time series data with seasonal features can be found in different fields, such as actuarial science, healthcare, economic, environment, and so on. They mostly display a seasonal template with periods, that repeat itself after a regular interval of time. The smallest time period for this event is called a seasonal period. Several factors such as weather and inherent attributes, can cause seasonal structures.

MSC(2020): 00A09, 54C40.

Keywords: Seasonality, α-fold zero modified geometric, Count time series, Delaporte distribution, INAR(1) models, Overdispersion.

Received: 15 January 2024, Accepted: 24 August 2024.

*Corresponding author.

In many applications, it is clear that the assumption of a continuous-valued range may not be appropriate. As an example, suppose that X_t be the number of individuals or events at time t, such that the outcome is integer-valued and hence discrete. These are time series arising from counting certain objects or events at specified times. The use of conventional autoregressive time series models such as $X_t = \rho \cdot X_{t-1} + \varepsilon_t$, for counting data is not recommended, because if the innovations of the model are counting and it is not necessary that X_t to be counted. Therefore, the idea of using a binomial thinning operator was proposed, which leads to the definition of “the integer-valued autoregressive model”, or abbreviated to the INAR model. The binomial thinning operator for the first time is defined in [27], as follows

$$\rho \circ X = \sum_{i=1}^{X} Y_i, \qquad X > 0, \tag{1.1}$$

and 0 otherwise, where the counting series $Y \coloneqq \{Y_i\}_{i\geq 1}$ is an independent and identical Bernoulli distributed with fixed success probability $\rho \in [0, 1]$. Also, non-negative integer-valued random variable X is independent of $Y_i' s$. Based on (1.1), for the first time, the INAR(1) model introduced in [3] and [23], as follows

$$X_t = \rho \circ X_{t-1} + \varepsilon_t, \qquad t \in Z, \tag{1.2}$$

where $0 \leq \rho < 1$, and $\{\varepsilon_t\}_{t\in Z}$ is a sequence i.i.d integer-valued random variables, called innovations and for each t, ε_t is independent of X_{t-s} for all $s \geq 1$, $E\left(\varepsilon_t\right) = \mu_\varepsilon$ and Var $\left(\varepsilon_t\right) = \sigma_\varepsilon^2$. From [3], for $\rho \in [0, 1)$ and $\rho = 1$, the process $\{X_t\}_{t\in Z}$ are (strictly) stationarity and non-stationarity, respectively. The autocorrelation function (ACF) of the model (1.2) is $\rho_k = Corr(X_t, X_{t-k}) = \rho^k$, for $k \geq 0$, that is, the same as ACF of AR(1), which takes the only non-negative values. Also, $\rho > 0$ or $\rho = 0$ implies that the observations of $\{X_t\}_{t\in Z}$ are dependent or independent. Various types of INAR(1) models have been introduced in the literature. For example, see the survey by [30], [19], [25] and [16]. The Binomial thinning INAR(1) with Poisson α-fold zero modified geometric innovations was studied in [26].

Overdispersion is an important concept in the analysis of discrete data. The Poisson integer-valued autoregressive process is not suitable for modeling overdispersed counts because the Poisson distribution is equidispersion. Various methods have been proposed to overcome it. One approach is to change the distribution of innovation. Another suggestion to deal with overdispersion is to change the type of thinning operator, which often changes the distribution of innovation. The third approach is to change the marginal distribution of the process. In [31, 32] it was shown that a reason for overdispersion is the presence of a positive correlation between the monitored events. As a result, several authors have proposed new models which considered overdispersed series [1, 2, 4, 8, 18, 14, 20, 21]. An INAR(1) process for modeling count time series with equidispersion, underdispersion and overdispersion was studied in [9]. Two INAR(1) processes with double Poisson, and generalized Poisson innovations for modeling non-negative integer-valued time series with equidispersion, underdispersion, or overdispersion were studied [6]. The use of mixed Poisson distributions is recommended to treat overdispersion in the count time series data. The class of overdispersed INAR(1) processes with marginals belonging to a general class of mixed Poisson distributions was studied in [5].

The cumulative Delaporte distribution is a discrete probability distribution that can be considered the convolution of a negative binomial distribution with a Poisson distribution. This distribution has three parameters α, β and λ, and introduced by Pierre Delaporte [12]. For the special value of parameters this distribution reduces to Poisson, Polya, or geometric distribution.

The survey of seasonal time series of counts can be seen in some papers. [17] used the Poisson INAR(1) model with explanatory variables through autoregressive coefficient and Poisson parameter, proposed by [10], and modeled the seasonality by selecting a mix of sine and cosine terms to be the explanatory variables. Also, [34] generalized the random coefficient INAR(1) model by using covariables in the parameters of innovations to explain seasonality. [35] introduced an INAR(1) model extended by adding a second independent non-stationary series with varying mean. [7] captured the seasonality by connecting the time series variables with the corresponding seasonal period based on the Poisson INAR(1) model. [11] Proposed a first order INAR process with seasonally varying autocorrelation parameters and intra-sensonality dependent shocks.

This first-order seasonal INAR process with seasonal period s $(\text{INAR}(1)_s)$ has a more straightforward model structure and less parameters to estimate. However, it is not appropriate for modeling seasonal time series with overdispersion. This suggested that to model this kind of time series, other thinning operator and distribution may be needed to be taken into consideration. [29] introduced a seasonal geometric INAR(1) process based on the negative binomial thinning operator. The main objective of this paper is to propose a new stationary seasonal integer-valued autoregressive model based on the binomial thinning operator with Delaporte marginal distribution. We denote this model by DELINAR$(1)_s$. The motivation for such a process arises from its potential in the modeling and analysis of non-negative integer-valued time series with a seasonal structure in the overdispersed situation.

The rest of the article is organized as follows. The model is introduced in Section 2 and the transition probabilities of the model based on marginal and innovation mass function is derived. Also, some of the statistical and conditional properties of the model are obtained. Section 3 is devoted to estimation methods for estimating the unknown parameters of the model. Section 4 discusses some theoretical results for point forecasts. Section 5 discusses some simulation results for the estimation methods. In Section 6, the model is applied to a real set of data. Finally, we conclude in Section 7.

2. The Model and Its Innovation Term

In this section, we study structural properties of $DELINAR(1)_s$ process, such as, the distributions of the marginal and innovation, mean and variance of these distributions, autocovariance function, conditional expectation and conditional variance of the marginal random variable, and transition probabilities.

2.1. The Delaporte distribution.

Definition 2.1. The Delaporte distribution is a discrete valued distribution with probability generating function (pgf)

$$G_X(t) = e^{-\lambda(1-t)} \left(\frac{1}{1+\beta(1-t)} \right)^{\alpha}, \tag{2.1}$$

where $|t| \leq 1, \lambda > 0, \alpha, \beta > 0$. The relation (2.1) shows that it is the convolution of a negative binomial and Poisson random variables (see, [12] and [33]). The probability mass function (pmf) corresponding to (2.1) is given by

$$f(x) = p(X = x) = \sum_{i=0}^{x} \frac{\Gamma(i+\alpha) e^{-\lambda} \lambda^{x-i} \beta^i}{\Gamma(\alpha) i! (1+\beta)^{\alpha+i} (x-i)!},$$

for $x = 0, 1, 2, \ldots$ and $\alpha, \beta, \lambda > 0$. This distribution is denoted by $\text{Del}(\lambda, \alpha, \beta)$.

Differentiating from (2.1) with respect to t, it is easy to show

$$\mu_x = E(X) = \lambda + \alpha\beta \qquad \text{and} \qquad \sigma_x^2 = Var(X) = \lambda + \alpha\beta(1+\beta).$$

Thus, the dispersion index, which is the variance-to-mean ratio, is given by

$$I_x = \frac{\sigma_X^2}{\mu_x} = 1 + \frac{\alpha\beta^2}{\lambda + \alpha\beta}.$$

It follows that Delaporte distribution shows overdispersion.

2.2. **The first-order seasonal non-negative Delaporte INAR model.**

Definition 2.2. A discrete-time non-negative integer-valued stochastic process $\{X_t\}$ is said to be a new seasonal Delaporte INAR process with seasonal period s ($DELINAR(1)_s$) based on the binomial thinning operator if it satisfies the following equation:

$$X_t = \rho \circ X_{t-s} + \varepsilon_t, \qquad t = 0, 1, 2, \ldots \tag{2.2}$$

where $s \in \mathbb{N}$ denotes the seasonal period and $\{X_t\}$ is a sequence of random variables with Delaporte distribution which previously introduced in Definition 2.1.

Note that when $s = 1$, this model degenerates into the Delaporate first-order INAR model, denoted by $DELINAR(1)$, so in this paper, we let $s \geq 2$. It can also be said that the ($DELINAR(1)_s$) process consists of s mutually independent INAR(1) processes with the same autoregressive coefficient ρ and the same innovation distribution. Let $X_t^{(r)} := X_{ts+r}, t \in \mathbb{N}_0, r = 1, 2, \ldots, s$, then it is easy to see that for all $r = 1, 2, \ldots, s$, the process $\{X_t^{(r)}\}$ satisfies the $DELINAR(1)$ model

$$X_t^{(r)} = \rho \circ X_{t-1}^{(r)} + \varepsilon_t^{(r)}$$

where the innovation sequence $\varepsilon_t^{(r)}$ is defined by $\varepsilon_t^{(r)} := \varepsilon_{ts+r}$. The independence of the stochastic processes $\{X_t^{(r)}\}$, clearly follows from the independences of the innovation sequences $\varepsilon_t^{(r)}$, and the counting processes involved in the thinning operators. The decomposition implies that the ($DELINAR(1)_s$) process $\{X_t\}$ is a so-called s-step Markov chain, that is, for all $t \geq s$,

$$P(X_t = x_t | X_{t-1} = x_{t-1}, \ldots, X_0 = x_0) = P(X_t = x_t | X_{t-s} = x_{t-s}),$$

for any $x_0, x_1, \ldots, x_n \in \mathbb{N}_0$.

The distribution of the innovation sequence $\{\varepsilon_t\}$ is given by the following proposition.

Proposition 2.3. *The random variable* ε_t *can be represented as* $\varepsilon_t = Y_1 + Y_2$ *where* $Y_1 \sim Po(\lambda(1-\rho))$ *and* Y_2 *are* α*-fold convolutions of zero-modified geometric distribution.*

Proof. Consider the first-order seasonal non-negative Delaporte INAR model $(DELINAR(1)_s)$ where X_t satisfies (2.2), If $G_X(z) = E(z^X)$ is the probability generating function of X, and $\Phi(z) = G(1-z)$ is the alternative probability generating function (apgf), the model defined in (2.2) in terms of apgf can be given as

$$\Phi_{X_t}(z) = \Phi_{X_{t-s}}(\rho z)\Phi_{\varepsilon_t}(z).$$

Under the stationarity of the process $\{X_t\}$, we have

$$\Phi_\varepsilon(z) = \frac{\Phi_X(z)}{\Phi_X(\rho z)} = e^{-\lambda(1-\rho)z}\left[\rho + (1-\rho)\frac{1}{1+\beta z}\right]^\alpha$$

Therefore, the innovation sequence $\{\varepsilon_t\}$ has convolution structure $\varepsilon_t = Y_1 + Y_2$ where $Y_1 \sim Po(\lambda(1-\rho))$ and Y_2 are α-fold convolutions of zero-modified geometric distribution. ☐

Using Proposition 2.3, since $Y_1 \sim Po(\lambda(1-\rho))$ and $Y_2 = \sum_{i=1}^{\alpha} Z_i$, where $Z_i \sim ZMG(p, p_0)$ and the fact that Y_1 and Y_2 are independent variables, we can obtained that the expectation and the variance of the random variable ε_t are

$$\mu_\varepsilon := E(\varepsilon_t) = (1-\rho)(\lambda + \alpha\beta) = (1-\rho)\mu_X,$$

and

$$\sigma_\varepsilon^2 := Var(\varepsilon_t) = (1-\rho)[\lambda + \alpha\beta(1 + (1+\rho)\beta)] = (1-\rho^2)\sigma_X^2 - \rho(1-\rho)\mu_X.$$

Also, we can derive the s-step transition probabilities of the $(DELINAR(1)_s)$ process by Proposition 1 and the s-step Markov property. It follows that

$$P(X_t = x_t | X_{t-s} = x_{t-s}) = P(\rho \circ X_{t-s} + \varepsilon_t = x_t | X_{t-s} = x_{t-s}).$$

Now, $B^\rho_{X_{t-s}} := \rho \circ X_{t-s} | X_{t-s} \sim Binomial(X_{t-s}, \rho)$, also ε_t is independent of X_{t-s} and is a convolution of two random variables with distribution of $P_{(\lambda,\rho)} \sim Po(\lambda(1-\rho))$ and α-fold convolutions of zero-modified geometric (ZMG) with parameters $p = \frac{1}{1+\beta}$ and p_0, where $p_0 = \frac{1+\rho\beta}{1+\beta}$ is probability mass at zero. Therefore, it can be written as

$$P(X_t = x_t | X_{t-s} = x_{t-s}) = P(B^\rho_{X_{t-s}} + P_{(\lambda,\rho)} + \alpha FZMG_{(p,p_0)} = x_t)$$

$$= \sum_{m=0}^{x_t} P(\alpha FZMG_{(p,p_0)} = m) \sum_{d=0}^{\min\{x_t - m, x_{t-s}\}} p(B^\rho_{X_{t-s}} = d) P(P_{(\lambda,\rho)} = x_t - m - d), \tag{2.3}$$

where,

$$P(P_{(\lambda,\rho)} = u) = \frac{e^{-\lambda(1-\rho)}(\lambda(1-\rho))^u}{u!}, \qquad u = 0, 1, 2, \dots$$
$$P(B^{\rho}_{X_{t-s}} = d) = \binom{X_{t-s}}{d}\rho^d(1-\rho)^{x_t - d}, \qquad d = 0, 1, 2, \dots, X_{t-s}$$

and $\alpha FZMG_{(p,p_0)} = \sum_{i=1}^{\alpha} Z_i$, such that

$$p(Z_i = k) = \begin{cases} p_0 & k = 0 \\ (1-\rho)\frac{\beta^k}{(1+\beta)^{k+1}} & k = 1, 2, \dots \end{cases}$$

where $p_0 = \rho + (1-\rho)\frac{1}{1+\beta}$ is probability mass at zero.

Proposition 2.4. *Suppose $X_{(k+h)s+j}, X_{ks+i}$ satisfy the process (2.2), and $h \in \mathbb{N}, k \in \mathbb{N}_0, i, j = 1, 2, \dots, s$. Then we have following results:*

(i) The conditional expectation of $X_{(k+h)s+j}$ given X_{ks+i} is

$$E(X_{(k+h)s+j}|X_{ks+i}) = \begin{cases} \mu_x & i \neq j \\ \rho^h X_{ks+i} + (1-\rho^h)\mu_x & i = j \end{cases}$$

and when $h \to \infty$, the above conditional expectation approaches to μ_x, which is the unconditional mean.

(ii) The conditional variance of $X_{(k+h)s+j}$ given X_{ks+i} is

$$Var(X_{(k+h)s+j}|X_{ks+i}) =$$
$$\begin{cases} \sigma^2_X, & i \neq j \\ \rho^h(1-\rho^h)X_{ks+i} + \frac{1-\rho^{2h}}{1-\rho^2}\sigma^2_\varepsilon + \rho(1-\rho)\left[\frac{1-\rho^{2(h-1)}}{1-\rho^2} - \frac{\rho^{h-1}(1-\rho^{h-1})}{1-\rho}\right]\mu_x, & i = j, \end{cases}$$

where $\sigma^2_\varepsilon := Var(\varepsilon_t) = (1-\rho)[\lambda + \alpha\beta(1 + (1+\rho)\beta)] = (1-\rho^2)\sigma^2_X - \rho(1-\rho)\mu_X$.

It is clear that when $h \to \infty$, $Var(X_{(k+h)s+j}|X_{ks+i}) \to \sigma^2_X$, which is the unconditional variance.

(iii) The autocovariance of the process is

$$Cov(X_{(k+h)s+j}, X_{ks+i}) = \begin{cases} 0, & i \neq j \\ \rho^h\sigma^2_x, & i = j, \end{cases}$$

Clearly, when $i = j$, the autocorrelation function $\rho(hs) = \rho^h$ decreases exponentially as $h \to \infty$. Note that μ_X and σ_X^2 are the unconditional mean and variance of the random variable X_n.

Proof. see the appendix. □

Proposition 2.5. *If $\rho \in [0, 1)$, the unique stationary marginal distribution of model (2.2) can be expressed in terms of the innovation process $\{\varepsilon_t\}$ as*

$$X_t \overset{d}{=} \sum_{k=0}^{\infty} \rho^k \circ \varepsilon_{t-ks} = \varepsilon_t + \sum_{k=1}^{\infty} \sum_{j=1}^{\varepsilon_{t-ks}} Y_{t,k,j}, \qquad t \in \mathbb{N}_0$$

where $\overset{d}{=}$ stands for equality in distribution. For all $t \in \mathbb{N}_0$, the Bernoulli variables $\{Y_{t,k,j,}\}_{k,j\geq 1}$ being mutually independent and independent of the innovation process with $E(Y_{t,k,j,}) = \rho^k$ for all $k, j \geq 1$. Also, the infinite series is understood as the limit in probability of the finite sum.

Proof. see the appendix. □

3. **Estimation methods**

Let us assume that we have n observations $x_1, x_2, \ldots, x_n$ from $DELINAR(1)_s$ process. In the $DELINAR(1)_s$ model we have four parameters. As we know, α takes non-negative integer values, so for simplicity of calculations, we assume that it is known, then after estimating other parameters of the model, we fit model with different values of α to choose the most optimal one. Therefore, three parameters ρ, λ and β have to be estimated. Three methods will be considered in this section, Yule-Walker method (YW), conditional least squares method (CLS), and conditional maximum likelihood method (CML).

3.1. **Yule-Walker estimation (YW).**

Let $X_1, \ldots, X_n$ be a sample of process (2.2). The sample mean is $\bar{X} = \frac{1}{n}\sum_{t=1}^{n} X_t$, the sample variance is $S^2 = \frac{1}{n-1}\sum_{t=1}^{n}(X_t - \bar{X})^2$, and the sample autocorrelation function with lag 1 is

$$\hat{\rho}_1 = \frac{\sum_{t=s+1}^{n}(X_t - \bar{X})(X_{t-s} - \bar{X})}{\sum_{t=1}^{n}(X_t - \bar{X})^2}$$

Thus the Yule-Walker (YW) estimate of ρ is given by

$$\hat{\rho}_{YW} = \frac{\sum_{t=s+1}^{n}(X_t - \bar{X})(X_{t-s} - \bar{X})}{\sum_{t=1}^{n}(X_t - \bar{X})^2}$$

and $\hat{\mu}_{YW} = \bar{X}$. Since $E(X_t) = \lambda + \alpha\beta$ and $Var(X_t) = \lambda + \alpha\beta(1+\beta)$, the Yule-Walker estimators of λ and β are

$$\hat{\beta}_{YW} = \sqrt{\frac{S^2 - \overline{X}}{\alpha}} \qquad and \qquad \hat{\lambda}_{YW} = \bar{X} - \alpha\sqrt{\frac{S^2 - \overline{X}}{\alpha}}.$$

3.2. **Conditional least squares estimation (CLS).** In this subsection, we use the conditional least squares method and obtain the estimators for the parameters ρ, μ and σ^2. We obtain the conditional least squares estimations in two steps [22]. First, we apply the ordinary conditional least squares for estimation of the parameters ρ and μ. The ordinary conditional least squares for the model is given by

$$\begin{aligned} S_{1CLS} &= \sum_{t=s+1}^{n} \mathrm{e}_{1t}^2 = \sum_{t=s+1}^{n} (X_t - E[X_t|X_{t-s}])^2 \\ &= \sum_{t=s+1}^{n} (X_t - \rho X_{t-s} - (1-\rho)(\lambda + \alpha\beta))^2. \end{aligned}$$

Therefore, by taking partial derivative with respect to ρ and μ, after some calculation, the conditional least squares estimators of the parameters ρ and μ are given by

$$\begin{aligned} \hat{\rho}_{CLS} &= \frac{(n-s)\sum_{t=s+1}^{n} X_t X_{t-s} - \sum_{t=s+1}^{n} X_t \sum_{t=s+1}^{n} X_{t-s}}{(n-s)\sum_{t=s+1}^{n} X_{t-s}^2 - (\sum_{t=s+1}^{n} X_{t-s})^2}, \\ \hat{\mu}_{cls} &= \frac{\sum_{t=s+1}^{n} X_t - \hat{\rho}_{cls}\sum_{t=s+1}^{n} X_{t-s}}{(n-s)(1-\hat{\rho}_{cls})}. \end{aligned}$$

For the estimation of the parameter σ^2, in step two, let $V_t = (X_t - E(X_t|X_{t-s}))^2 = (X_t - \rho X_{t-s} - (1-\rho)\mu)^2$ be a random variable. It is easy to show that $E(V_t|X_{t-s}) = Var(X_t|X_{t-s}) = \rho(1-\rho)X_{t-s} + (1-\rho^2)\sigma^2 - \rho(1-\rho)\mu$. Using V_t and its conditional expectation, we have a new sum of squares that is given by

$$S_{2CLS}(\sigma^2) = \sum_{t=s+1}^{n} (V_t - E[V_t|X_{t-s}])^2 = \sum_{t=s+1}^{n} (V_t - \rho(1-\rho)(X_{t-s} - \mu) - (1-\rho^2)\sigma^2)^2.$$

By replacing the estimators of ρ and μ from previous step in $S_{2CLS}(\sigma^2)$ and minimizing $S_{2CLS}(\sigma^2)$ with respect to σ^2, we have

$$\sigma_{cls}^2 = \frac{\sum_{t=s+1}^{n} (X_t - \hat{\rho}_{CLS}X_{t-s} - (1-\hat{\rho}_{CLS})\,\hat{\mu}_{cls})^2 - \hat{\rho}_{CLS}\,(1-\hat{\rho}_{CLS}) \sum_{t=s+1}^{n} (X_{t-s} - \hat{\mu}_{cls})}{(n-s)(1-\hat{\rho}_{cls}^2)}$$

At last, since under the stationary condition, we have $\mu = E(X_t) = \lambda + \alpha\beta$ and $\sigma^2 = Var\,(X_t) = \;\lambda + \alpha\beta\,(1+\beta)$, the conditional least squares estimators for the parameters λ, β are given by

$$\hat{\beta}_{cls} = \sqrt{\frac{\hat{\sigma}^2_{cls} - \hat{\mu}_{cls}}{\alpha}}, \qquad \hat{\lambda}_{cls} = \hat{\mu}_{cls} - \alpha\hat{\beta}_{cls}.$$

3.3. **Conditional maximum likelihood estimation (CML).** Suppose $X_1, X_2, \ldots, X_n$ be a random sample from a stationary $\text{DELINAR}(1)_s$ process with parameters ρ, λ and β. The conditional log-likelihood function is given by

$$CL(\rho, \lambda, \beta) = \sum_{t=s+1}^{n} \log P(X_t = j | X_{t-s} = i)$$

where $P(X_t = j|X_{t-s} = i)$ is defined in (2.3). The conditional maximum likelihood estimators are obtained by maximizing $CL(\rho, \lambda, \beta)$. Since derivative of $CL(\rho, \lambda, \beta)$ is a nonlinear function, the maximum likelihood estimate of parameters must be computed using numerical methods.

4. Forecasting

In this section, forecasting a future value X_{n+h}, $h \in \mathbb{N}$ based on the past information up to time n is expressed. The distribution of X_{n+h} based on the definition of the $\text{DELINAR}(1)_s$ model can be represent as

$$X_{n+h} \overset{d}{=} \rho^k \circ X_{n+h-qs} + \sum_{j=0}^{q-1} \rho^j \circ \varepsilon_{n+h-js} \tag{4.1}$$

where $q := \left\lceil \frac{h}{s} \right\rceil$, with $\lceil x \rceil := \min\{n \in \mathbb{N} | x \leq n\}$. According to equation (4.1), the distribution of h-step ahead forecasting has a very complicated form, so helping conditional expectation $E(X_{n+h}|\mathcal{F}_n)$, we get the h-step ahead forecasting. $E(X_{n+h}|\mathcal{F}_n)$ given by

$$\begin{aligned}
E(X_{n+h}|\mathcal{F}_n) &= E(\rho^q \circ X_{n+h-qs} + \sum_{j=0}^{q-1} \rho^j \circ \varepsilon_{n+h-js} | \mathcal{F}_n) \\
&= \rho^q X_{n+h-qs} + \sum_{j=0}^{q-1} \rho^j (1-\rho)\mu_x \\
&= \rho^q (X_{n+h-qs} - \mu) + \mu,
\end{aligned}$$

where $\mu_x = \lambda + a\beta$ and $\mathcal{F}_n$ is the information set up to time n.

Some properties of the h-step ahead conditional expectation for DELINAR$(1)_s$ model are given in the following proposition:

Proposition 4.1. *Let $\{X_n\}$ be a stationary $DELINAR(1)_s$ process and $n, h \in N$. Then:*

(i) $E(X_{n+h}|\mathcal{F}_n) = \rho^q(X_{n+h-qs} - \mu_x) + \mu_x$

(ii) $Var(X_{n+h}|\mathcal{F}_n) = \rho^q(1-\rho^q)X_{n+h-qs} + \frac{1-\rho^{2q}}{1-\rho^2}\sigma_\varepsilon^2 + \left[(1-\rho^q) - \left(\frac{1-\rho^2 q}{1+\rho}\right)\right]\mu_x$

(iii) $\lim_{h\to\infty} E(X_{n+h}|\mathcal{F}_n) = \mu_x$

(iv) $\lim_{h\to\infty} Var(X_{n+h}|\mathcal{F}_n) = \sigma_x^2$

where $h \in \mathbb{N}$ and $q \coloneqq \lceil \frac{h}{s} \rceil$.

So, a forecast $\hat{X}_{n+h}, h \in \mathbb{N}$, based on the sample $X_1, X_2, \ldots, X_n$, is obtained by

$$\hat{X}_{n+h} = \hat{\rho}^q(X_{n+h-qs} - \hat{\mu}_x) + \hat{\mu}_x$$

where $\hat{\rho}, \hat{\mu}_x$ are estimators for ρ and μ, respectively.

Proof. Let $h \in \mathbb{N}$ and $q\colon = \lceil \frac{h}{s} \rceil$.

(i) Using (4.1), $E(\rho \circ X|X) = \rho X$, where X is a non-negative integer-valued random variable, and the fact that ε_{n+h-js} is independent of the $\mathcal{F}_n$ for all $j = 0, 1, \ldots, q-1$, $E(X_{n+h}|\mathcal{F}_n)$ is computed as follows:

$$\begin{aligned} E(X_{n+h}|\mathcal{F}_n) &= E\left(\rho^q \circ X_{n+h-qs} + \sum_{j=0}^{q-1} \rho^j \circ \varepsilon_{n+h-js}|\mathcal{F}_n\right) \\ &= \rho^q X_{n+h-qs} + \sum_{j=0}^{q-1} \rho^j (1-\rho)\mu_x \\ &= \rho^q(X_{n+h-qs} - \mu_x) + \mu_x, \end{aligned}$$

where $\mu_x = \lambda + \alpha\beta$.

(ii) Using (4.1), $Var(\rho \circ X|X) = \rho(1-\rho)X$, where X is a non-negative integer-valued random variable, and the fact that ε_{n+h-js} is independent of the $\mathcal{F}_n$ for all $j = 0, 1, \ldots, q-1$, $Var(X_{n+h}|\mathcal{F}_n)$ is computed as follows

$$Var(X_{n+h}|\mathcal{F}_n) = Var\left(\rho^q \circ X_{n+h-qs} + \sum_{j=0}^{q-1} \rho^j \circ \varepsilon_{n+h-js}|\mathcal{F}_n\right)$$

$$= Var\left(\rho^q \circ X_{n+h-qs} \middle| X_{n+h-qs}\right) + Var\Big(\sum_{j=0}^{q-1} \rho^j \circ \varepsilon_{n+h-js}\Big)$$

$$= \rho^q(1-\rho^q)\, X_{n+h-qs} + \sum_{j=0}^{q-1} Var(\rho^j \circ \varepsilon_{n+h-js})$$

$$= \rho^q(1-\rho^q)\, X_{n+h-qs} + \frac{1-\rho^{2q}}{1-\rho^2}\sigma_\varepsilon^2 + \Big[(1-\rho^q) - \Big(\frac{1-\rho^{2q}}{1+\rho}\Big)\Big]\mu_x.$$

Because we have

$$\begin{aligned} Var(\rho^j \circ \varepsilon_{n+h-js}) &= E\left(\rho^j \circ \varepsilon_{n+h-js}\right)^2 - E^2(\rho^j \circ \varepsilon_{n+h-js}) \\ &= \rho^{2j}E(\varepsilon^2) + \rho^j\left(1-\rho^j\right)E(\varepsilon) - \rho^{2j}E^2(\varepsilon) \\ &= \rho^{2j}\left(E(\varepsilon^2) - E^2(\varepsilon)\right) + \rho^j\left(1-\rho^j\right)E(\varepsilon) \\ &= \rho^{2j}\sigma_\varepsilon^2 + \rho^j(1-\rho^j)E(\varepsilon), \end{aligned}$$

where

$E(\varepsilon) = (1-\rho)\mu_x,$
$\sigma_\varepsilon^2 = Var(\varepsilon) = (1-\rho^2)\ \sigma_X^2 - \rho(1-\rho)\mu_X$ and $\sum_{j=0}^{q-1}\rho^{2j} = \frac{1-\rho^{2q}}{1-\rho^2}$.

(iii) and (iv) using the results given in (i) and (ii), the following limits are obtained:

$$\lim_{h\to\infty} E(X_{n+h}|\mathcal{F}_n) = \lim_{h\to\infty}[q(Xn+h-qs-\mu_x)+\mu_x] = \mu_x,$$

$$\begin{aligned} \lim_{h\to\infty} Var(X_{n+h}|\mathcal{F}_n) &= \lim_{h\to\infty}\Big[\rho^q(1-\rho^q)\, X_{n+h-qs} + \frac{1-\rho^{2q}}{1-\rho^2}\sigma_\varepsilon^2 + \Big[(1-\rho^q) - \Big(\frac{1-\rho^{2q}}{1+\rho}\Big)\Big]\mu_x\Big] \\ &= \sigma_x^2. \end{aligned}$$

□

5. **Monte Carlo simulation study**

At the beginning of this section, we simulate a sample of the new process with seasonal period s = 12 and $\rho = 0.8$, $\beta = 3$ and $\lambda = 1$. Figure 1 shows the sample path and the ACF of model as can be seen, the ACF of model is zero except at lags that are multiples of s. Also, the ACF, $\rho(k)$, decays exponentially with lag k All simulations have been performed in R programming. We generate the sample $n \in \{200, 400, 800\}$ from the DELINAR$(1)_s$ process with $s = 12$, and the number of replications for each n is 1000. In this simulation, we set (a) $(\rho, \beta, \lambda) = (0.2, 3, 1.5)$, (b) $(\rho, \beta, \lambda) = (0.5, 2, 1)$

and (c) $(\rho, \beta, \lambda) = (0.8, 1.5, 4)$. Tables 1, 2 and 3 show the results of this simulation in terms of bias and mean square error (MSE) of the estimators obtained from the three procedure method estimation. These tables show that the values of the bias and mean square error of the estimates of the parameters go to zero as the sample size increases for all cases. As can be seen from the tables, the conditional least square and the Yule-Walker methods show similar mean square behaviors. However, the conditional maximum likelihood estimators have the best implementation on bias and MSE compared with the Yule-Walker and the conditional least square estimators. It is because that both bias and MSE for the conditional maximum likelihood estimators are smaller than those for the other methods. Since the bias of $\hat{\rho}$ and $\hat{\beta}$ are negative for all estimation methods, so they tend to underestimate of the parameters, and since the bias of the estimators of $\hat{\lambda}$ is positive for all estimation methods, so it tends to overestimate the parameter.

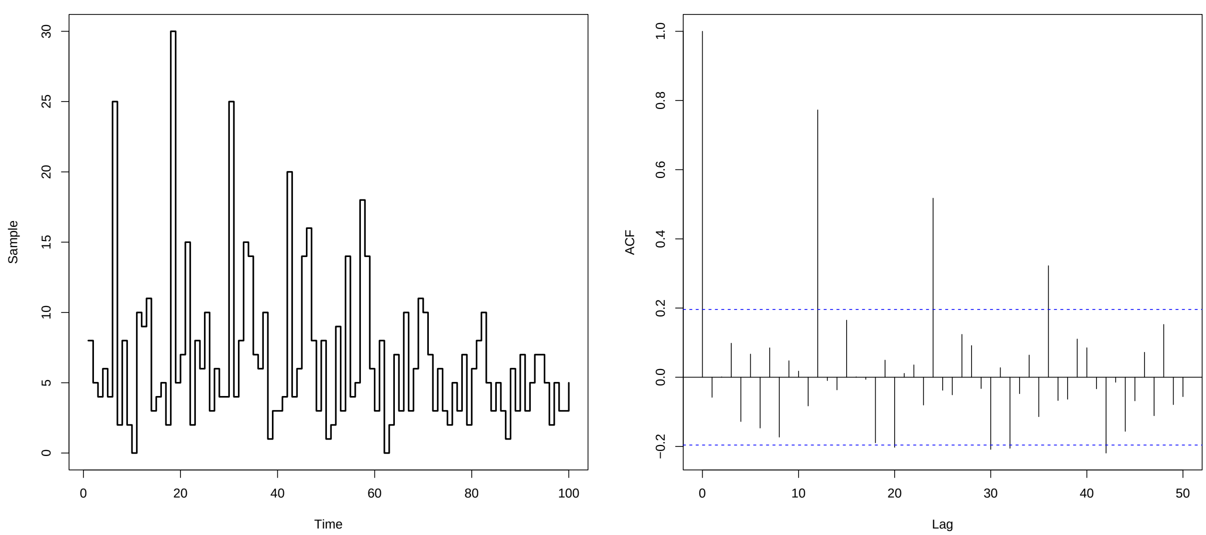

FIGURE 1. sample path of $DELINAR(1)_s$ process and its sample ACF

TABLE 1. Results of simulation: Bias and MSE (in parantheses) of estimators of parameters for $(\rho, \beta, \lambda) = (0.2, 3, 1.5)$.

n	$\hat{\rho}_{YW}$	$\hat{\rho}_{CLS}$	$\hat{\rho}_{CML}$	$\hat{\beta}_{YW}$	$\hat{\beta}_{CLS}$	$\hat{\beta}_{CML}$	$\hat{\lambda}_{YW}$	$\hat{\lambda}_{CLS}$	$\hat{\lambda}_{CML}$
200	−0.0142	−0.0023	0.0008	−0.0106	−0.0320	−0.0159	0.0075	0.0492	0.0176
	(0.0047)	(0.0050)	(0.0027)	(0.0956)	(0.1013)	(0.0766)	(0.2685)	(0.2804)	(0.1708)
400	−0.0074	−0.0013	0.0000	−0.0051	−0.0131	−0.0082	−0.0003	0.0155	0.0060
	(0.0025)	(0.0026)	(0.0013)	(0.0474)	(0.0498)	(0.0374)	(0.1344)	(0.1424)	(0.0870)
800	−0.0059	−0.0029	−0.0016	−0.0057	−0.0105	−0.0067	0.0095	0.0196	0.0119
	(0.0013)	(0.0013)	(0.0007)	(0.0262)	(0.0268)	(0.0189)	(0.0759)	(0.0774)	(0.0465)

TABLE 2. Results of simulation: Bias and MSE (in parantheses) of estimators of parameters for $(\rho, \beta, \lambda) = (0.5, 2, 1)$.

n	$\hat{\rho}_{YW}$	$\hat{\rho}_{CLS}$	$\hat{\rho}_{CML}$	$\hat{\beta}_{YW}$	$\hat{\beta}_{CLS}$	$\hat{\beta}_{CML}$	$\hat{\lambda}_{YW}$	$\hat{\lambda}_{CLS}$	$\hat{\lambda}_{CML}$
200	−0.0385	−0.0086	−0.0029	−0.0235	−0.0461	−0.0118	0.0460	0.0953	0.0272
	(0.0054)	(0.0044)	(0.0019)	(0.0808)	(0.0908)	(0.0660)	(0.2459)	(0.2779)	(0.1555)
400	−0.0231	−0.0081	−0.0034	−0.0278	−0.0367	−0.0163	0.0453	0.0645	0.0254
	(0.0026)	(0.0022)	(0.0008)	(0.0426)	(0.0451)	(0.0314)	(0.1278)	(0.1353)	(0.0725)
800	−0.0109	−0.0037	−0.0015	−0.0130	−0.0177	−0.0068	0.0142	0.0233	0.0017
	(0.0012)	(0.0011)	(0.0004)	(0.0209)	(0.0214)	(0.0147)	(0.0596)	(0.0608)	(0.0332)

TABLE 3. Results of simulation: Bias and MSE (in parantheses) of estimators of parameters for $(\rho, \beta, \lambda) = (0.8, 1.5, 4)$.

n	$\hat{\rho}_{YW}$	$\hat{\rho}_{CLS}$	$\hat{\rho}_{CML}$	$\hat{\beta}_{YW}$	$\hat{\beta}_{CLS}$	$\hat{\beta}_{CML}$	$\hat{\lambda}_{YW}$	$\hat{\lambda}_{CLS}$	$\hat{\lambda}_{CML}$
200	−0.0585	−0.0103	−0.0030	−0.0654	−0.1241	−0.0159	0.1532	0.2602	0.0461
	(0.0057)	(0.0022)	(0.0005)	(0.1936)	(0.2676)	(0.1088)	(0.8151)	(1.2005)	(0.4488)
400	−0.0326	−0.0081	−0.0021	−0.0594	−0.0888	−0.0111	0.1120	0.1768	0.0180
	(0.0021)	(0.0010)	(0.0002)	(0.1165)	(0.1334)	(0.0555)	(0.4940)	(0.5645)	(0.2416)
800	−0.0169	−0.0046	−0.0009	−0.0402	−0.0502	−0.0100	0.0907	0.1130	0.0318
	(0.0008)	(0.0005)	(0.0001)	(0.0625)	(0.0676)	(0.0269)	(0.2486)	(0.2673)	(0.1130)

6. Real data example

The model proposed in Section 2, is now used to model and forecast the counts of claims of short-term disability benefits by month. The data are included is tsinteger package and can be taken from `http://rdrr.io>github>portriaota>tsinteger`. This data consists the male claimants, between 35 and 54 ages, work in the logging industry and reported their claim to the Richmond, BC Workers Compensation Board. In the data set, only claimants whose injuries were due to cuts and lacerations were consisted. The 120 observation was gathered from January 1985 to December 1994. [15], [35] and [7] were formerly analyzed this data set.

The sample mean and the sample variance of data are 6.13 and 11.80, respectively. The Fisher index of dispersion of the data set is 1.92. We conclude that the data set has overdispersed property. In Figure 2, the sample path of time series, the sample autocorrelation function (ACF) and partial autocorrelation function (PACF) are exhibited. The geometrical decrease pattern with a seasonal period of 12 can be seen in ACF plot and detects that there is a time series with serial correlations behaviors. According to the plots, it is clear that the data has overdispersion and seasonality property, so we are encouraged to choose our DELINAR$(1)_{12}$ model as a candidate model, also we will consider the PINAR(1), DElINAR(1) and INAR$(1)_{12}$ models for comparison. The data set will be divided into two parts. We use the first 110 observations to model the series, also for forecasting purposes, the last 10 observations are considered.

Alpha takes non-negative integer values, so we fitted model to the data with values of 1, 2 and 3 for α by CML approach to select the best one according to the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). The results are presented in Tables 4. This table also shows the CML estimates (with MSE in parentheses) of the parameters of the $DELINAR(1)_{12}$ process. According to sensitivity analysis, it seems that the AIC increases slightly as the α value increases. So we considered $\alpha = 1$ for model.

Table 5, the CML estimates (with corresponding mean square errors in parentheses) of parameters for fitting the model DELINAR$(1)_{12}$ and the other models, AIC and BIC are given. For the model DELINAR$(1)_{12}$ the values of the AIC and BIC are smaller than the values of other models.

TABLE 4. Sensitivity analysis to select best value for α.

α	ρ	β	λ	AIC	BIC
1	0.2285(0.0104)	2.7345(0.2032)	3.5575(0.4816)	500.21	508.31
2	0.2299(0.0059)	1.9001(0.2549)	2.7492(0.6240)	500.54	508.64
3	0.2097(0.0076)	1.5266(0.0238)	1.9548(0.0315)	500.91	509.01

Then, the suitable model by CML estimation is

$$X_t = 0.2285 \circ X_{t-12} + \varepsilon_t$$

and X_t follows a $Del(3.5575, 1, 2.7345)$ and ε_t is a convolution of two random variables with distribution of $Poisson(2.7446)$ and zero-modified geometric with parameters $p = 0.2677$, and $p_0 = 0.4350$. In Figure 3, the observed, fitted and forecasting values are shown, with black, red and blue color lines, respectively. Figure 4 shows ACF and PACF of residuals, as can be seen, serial correlations are still observed in the residuals.

TABLE 5. Estimated parameters (MSE in parentheses), AIC and BIC.

$Model$	$CML estimate$	$CLS estimate$	AIC	BIC
$PINAR(1)$	$\hat{\rho} = 0.5705(0.0005)$	$\hat{\rho} = 0.5651(0.0075)$	545.80	551.20
	$\hat{\lambda} = 2.7059(0.0188)$	$\hat{\lambda} = 2.7356(0.3199)$		
$INAR(1)_{12}$	$\hat{\rho} = 0.2459(0.004)$	$\hat{\rho} = 0.2667(0.0098)$	532.09	537.49
	$\hat{\lambda} = 4.6769(0.0411)$	$\hat{\lambda} = 4.5389(0.4201)$		
$DELINAR(1)$	$\hat{\rho} = 0.4436(0.0016)$	$\hat{\rho} = 0.5651(0.0070)$	527.08	535.18
	$\hat{\beta} = 1.2768(0.0077)$	$\hat{\beta} = 1.6831(0.1000)$		
	$\hat{\lambda} = 3.6945(0.1148)$	$\hat{\lambda} = 2.9253(0.3223)$		
$DELINAR(1)_{12}$	$\hat{\rho} = 0.2285(0.0104)$	$\hat{\rho} = 0.2667(0.0100)$	500.21	508.31
	$\hat{\beta} = 2.7345(0.2032)$	$\hat{\beta} = 2.5089(0.3197)$		
	$\hat{\lambda} = 3.5575(0.4816)$	$\hat{\lambda} = 3.6811(0.2911)$		

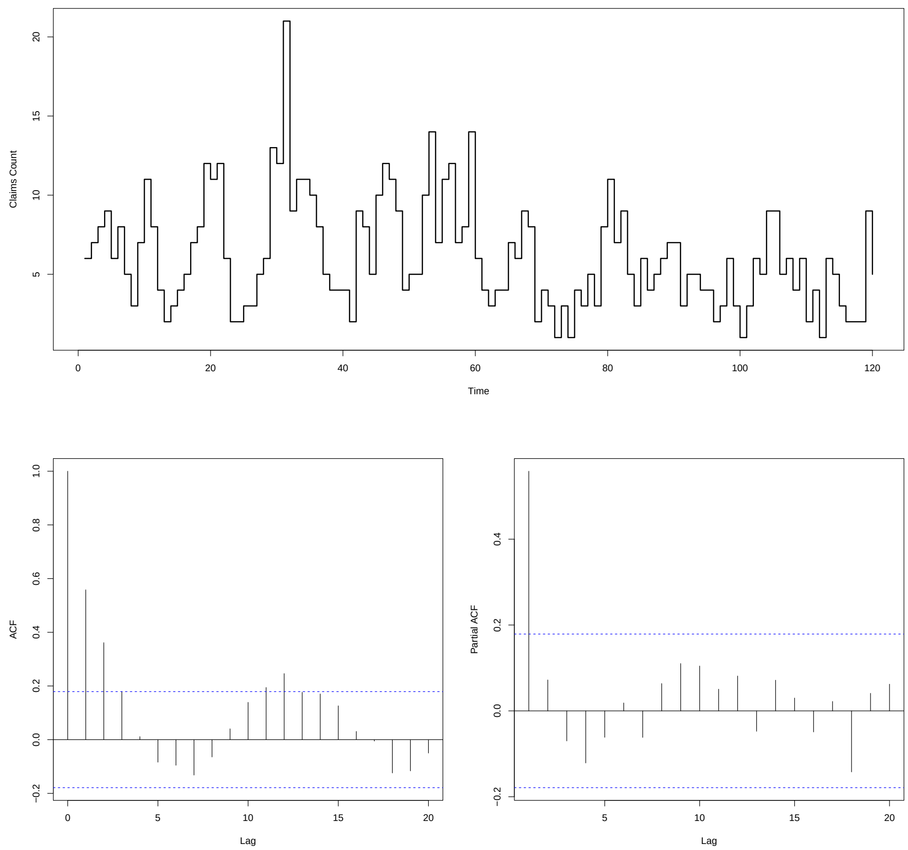

FIGURE 2. The time series, ACF and PACF of the claims series from 1985 to 1994.

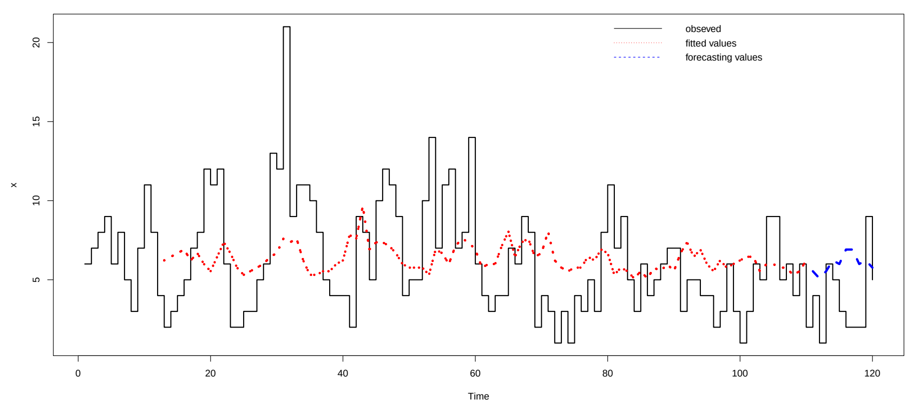

FIGURE 3. The plot of observed, fitted and forecasting values.

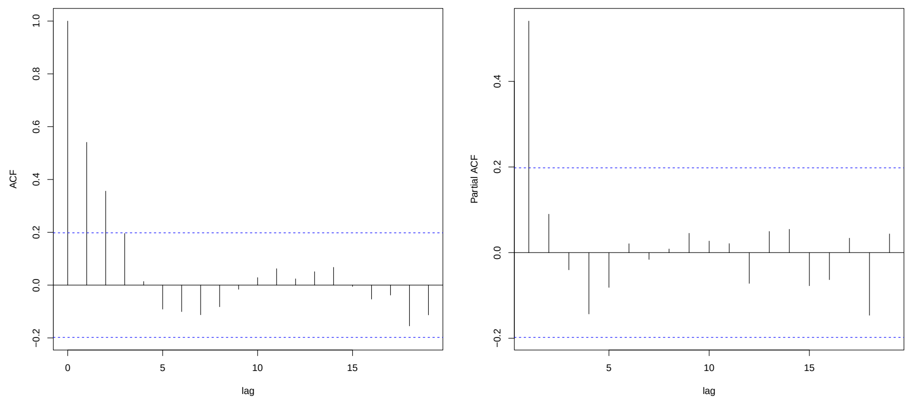

Figure 4. The ACF and PACF of residuals.

7. Conclusions

In this paper, we proposed a new seasonal Delaporte integer-valued autoregressive model with the binomial thinning operator. This model is appropriate for the seasonal and overdispersed data. We derive main properties of this model, consider three estimators for the model parameters (YW, CLS and CML) and compare these methods via simulation study. The results show that the Yule-Walker and conditional least square methods have similar performances, but the conditional maximum likelihood estimators are better than others. Therefore, we propose conditional maximum likelihood estimators for the parameters of the DELINAR$(1)_{12}$. We also discuss the forecasting values of the model by conditional expectation method. Finally, we fitted the model to real data set. The PINAR(1), INAR$(1)_{12}$ and DELINAR(1) models are also used to fit the same data for comparison. The result shows that based on the AIC, BIC criterion, our model is better than compared to the other INAR models. As part of future research, it would be interested to extend the model to autoregressive order $p > 1$. Also, according to the ACF of residuals shown in Figure 4, this model does not capture serial correlation. Therefore, the suggestion of a new model that can capture the both seasonal and serial correlation is necessary.

appendix

Proof of Proposition 2.4. Let $X_{(k+h)s+j} := X_{k+h}^{(j)}$ and $X_{ks+i} := X_k^{(i)}$. Then we have

(i): When $i \neq j$, $X_{k+h}^{(j)}$ and $X_k^{(i)}$ are mutually independent, so the conditional expectation equals to the unconditional expectation μ. When $i = j$, we can derive that

$$\begin{aligned}
E\left(X_{(k+h)s+j} \mid X_{ks+i}\right) &= E\left(X_{k+h}^{(j)} \mid X_k^{(j)}\right) \\
&= E\left(\rho \circ X_{k+h-1}^{(j)} + \varepsilon_{k+h}^{(j)} \mid X_k^{(j)}\right) \\
&= \rho E\left(X_{k+h-1}^{(j)} \mid X_k^{(j)}\right) + E\left(\varepsilon_{k+h}^{(j)}\right) \\
&= \rho^2 E\left(X_{k+h-1}^{(j)} \mid X_k^{(j)}\right) + (\rho+1)[(1-\rho)\mu] \\
&= \cdots \\
&= \rho^h E\left(X_k^{(j)} \mid X_k^{(j)}\right) + \left(\rho^{h-1} + \cdots + \rho + 1\right)[(1-\rho)\mu] \\
&= \rho^h X_k^{(j)} + \frac{1-\rho^h}{1-\rho}[(1-\rho)\mu] \\
&= \rho^h X_k^{(j)} + \left(1-\rho^h\right)\mu.
\end{aligned}$$

From the result, we can prove that $\lim_{h\to\infty} E(X_{(k+h)s+j} | X_{ks+i}) = \mu$.

(ii): When $i \neq j$, $X_{k+h}^{(j)}$ and $X_k^{(i)}$ are mutually independent, so the conditional variance equals to the unconditional variance . When $i = j$, we can derive that

$$\begin{aligned}
\operatorname{Var}\left(X_{(k+h)s+j} \mid X_{ks+i}\right) =& \operatorname{Var}\left(X_{k+h}^{(j)} \mid X_k^{(j)}\right) \\
=& \operatorname{Var}\left(\rho \circ X_{k+h-1}^{(j)} + \varepsilon_{k+h}^{(j)} \mid X_k^{(j)}\right) \\
=& \operatorname{Var}\left(\rho \circ X_{k+h-1}^{(j)} \mid X_k^{(j)}\right) + \operatorname{Var}\left(\varepsilon_{k+h}^{(j)} \mid X_k^{(j)}\right) \\
=& E\left(\left(\rho \circ X_{k+h-1}^{(j)}\right)^2 \mid X_k^{(j)}\right) - \left[E\left(\rho \circ X_{k+h-1}^{(j)} \mid X_k^{(j)}\right)\right]^2 \\
& + \operatorname{Var}\left(\varepsilon_{k+h}^{(j)}\right) \\
=& \rho^2 E\left[\left(\rho \circ X_{k+h-1}^{(j)}\right)^2 \mid X_k^{(j)}\right] + \rho(1-\rho) E\left(X_{k+h-1}^{(j)} \mid X_k^{(j)}\right) \\
& - \rho^2\left[E\left(X_{k+h-1}^{(j)} \mid X_k^{(j)}\right)\right]^2 + \sigma_\varepsilon^2 \\
=& \cdots \\
=& \rho^{2h} \operatorname{Var}\left(X_k^{(j)} \mid X_k^{(j)}\right) + \rho(1-\rho)\left[\rho^{2(h-1)} E\left(X_k^{(j)} \mid X_k^{(j)}\right)\right.
\end{aligned}$$

$$+\cdots+E\left(X_{k+h-1}^{(j)} \mid X_k^{(j)}\right)\Big]+\left(\rho^{2(h-1)}+\cdots+\rho^2+1\right)\sigma_{\varepsilon}^2,$$

where $Var(X_k^{(j)}|X_k^{(j)})=0$ and $(\rho^{2(h-1)}+\cdot+\rho^2+1)\sigma_\varepsilon^2=\frac{1-\rho^{2h}}{1-\rho^2}\sigma_\varepsilon^2$, also the second item of the above formula given by

$$\begin{aligned}
&\rho(1-\rho)\left[\rho^{2(h-1)}E\left(X_k^{(j)} \mid X_k^{(j)}\right)+\rho^{2(h-2)}E\left(X_{k+1}^{(j)} \mid X_k^{(j)}\right)+\cdots+\right.\\
&\left.\rho^2E\left(X_{k+h-2}^{(j)} \mid X_k^{(j)}\right)+E\left(X_{k+h-1}^{(j)} \mid X_k^{(j)}\right)\right]\\
&=\rho(1-\rho)\left\{\rho^{2(h-1)}X_k^{(j)}+\rho^{2(h-2)}\left[\rho X_k^{(j)}+\frac{1-\rho}{1-\rho}\mu_\varepsilon\right]\right.\\
&\quad\left.+\cdots+\rho^2\left[\rho^{h-2}X_k^{(j)}+\frac{1-\rho^{h-2}}{1-\rho}\mu_\varepsilon\right]+\left[\rho^{h-1}X_k^{(j)}+\frac{1-\rho^{h-1}}{1-\rho}\mu_\varepsilon\right]\right\}\\
&=\rho(1-\rho)\left\{\left[\rho^{2h-2}+\rho^{2h-3}+\cdots+\rho^{h-1}\right]X_k^{(j)}\right.\\
&\quad\left.+\left[\rho^{2(h-2)}(1-\rho)+\cdots+\rho^2\left(1-\rho^{h-2}\right)+\left(1-\rho^{h-1}\right)\right]\mu\right\}\\
&=\rho(1-\rho)\left\{\frac{\rho^{h-1}\left(1-\rho^h\right)}{1-\rho}X_k^{(j)}+\left[\frac{(1-\rho^{2(h-1)})}{(1-\rho^2}-\frac{\rho^{h-1}\left(1-\rho^{h-1}\right)}{1-\rho}\right]\mu\right\}.
\end{aligned}$$

Hence, when $i=j$, we have:

$$\begin{aligned}
Var(X_{(k+h)s+j}|X_{ks+i})=&\frac{1-\rho^{2h}}{1-\rho^2}\sigma_\varepsilon^2+\rho^h(1-\rho^h)X_k^j\\
&+\rho(1-\rho)\Big[\frac{1-\rho^{2(h-1)}}{1-\rho^2}-\frac{\rho^{h-1}(1-\rho^{h-1})}{1-\rho}\Big]\mu,
\end{aligned}$$

where $\sigma_\varepsilon^2=Var(\varepsilon_t)=(1-\rho)[\lambda+\alpha\beta(1+(1+\rho)\beta)]=(1-\rho^2)\sigma_X^2-\rho(1-\rho)\mu_X$. It is clear that the following limit is obtained: $\lim_{h\to\infty}Var(X_{(k+h)s+j}|X_{ks+i})=\sigma_X^2$.

(iii): When $i\neq j$, $X_{k+h}^{(j)}$ and $X_k^{(i)}$ are mutually independent, so the covariance equals to 0. When $i=j$, we can derive that

$$\begin{aligned}
&\operatorname{Cov}\left(X_{(k+h)k+j},X_{ks+i}\right)=\operatorname{Cov}\left(X_{k+h}^{(j)},X_k^{(j)}\right)=E\left(X_{k+h}^{(j)}X_k^{(j)}\right)-E\left(X_{k+h}^{(j)}\right)\left(X_k^{(j)}\right)\\
&=E\left[\left(\rho\circ X_{k+h-1}^{(j)}+\varepsilon_{k+h}^{(j)}\right)X_k^{(j)}\right]-E\left(\rho\circ X_{k+h-1}^{(j)}+\varepsilon_{k+h}^{(j)}\right)E\left(X_k^{(j)}\right)\\
&=E\left[\left(\sum_{p=1}^{x_{k+h-1}^{(j)}}Y_p\right)X_k^{(j)}\right]+E\left(\varepsilon_{k+h}^{(j)}\right)E\left(X_k^{(j)}\right)-\rho E\left(X_{k+h-1}^{(j)}\right)E\left(X_k^{(j)}\right)\\
&\quad-E\left(\varepsilon_{k+h}^{(j)}\right)E\left(X_k^{(j)}\right)
\end{aligned}$$

$$= E\left\{ E\left[\sum_{p+1}^{x_{k-1}^{(j)}} Y_p \mid X_{k+h-1}^{(j)} \right] X_k^{(j)} \right\} - \rho E\left(X_{k+h-1}^{(j)}\right) E\left(X_k^{(j)}\right)$$

$$= E\left(X_{k+h-1}^{(j)} X_k^{(j)} E\left(Y_p\right)\right) - \rho E\left(X_{k+h-1}^{(j)}\right) E\left(X_k^{(j)}\right)$$

$$= \rho \operatorname{Cov}\left(X_{k+h-1}^{(j)}, X_k^{(j)}\right)$$

$$\cdots$$

$$= \rho^h \operatorname{Cov}\left(X_k^{(j)}, X_k^{(j)}\right)$$

$$= \rho^h \operatorname{Var}\left(X_k^{(j)}\right)$$

$$= \rho^h \sigma_X^2.$$

Using the result, we can prove that the ACF of the process $\rho(hs) = \rho^h$ decays exponentially as $h \to \infty$.

Proof of Proposition 2.5. Define the random variable $Z_{t,n}$ as

$$Z_{t,n} = \sum_{k=0}^{n} \rho^k \circ \varepsilon_{t-ks} = \varepsilon_t + \sum_{k=1}^{n} \sum_{j=1}^{\varepsilon_{t-ks}} Y_{t,k,j}$$

Where the Bernoulli variables $\{Y_{t,k,j,}\}_{k,j\geq 1}$ being mutually independent and independent of the innovation process with $E(Y_{t,k,j,}) = \rho^k$ for all $k, j \geq 1$ and $t \in \mathbb{N}_0$. If μ_ε and σ_ε^2 represent, respectively, the mean and variance of the innovation sequence, by using properties (iv) and (v) of the thinning operator, see [28], we obtain for all $0 < n < m$,

$$E\Big[\sum_{k=n}^{m} \rho^k \circ \varepsilon_{t-ks}\Big] = \sum_{k=n}^{m} \rho^k \mu_\varepsilon \leq \frac{\rho^n}{1-\rho} \mu_\varepsilon$$

$$Var\Big[\sum_{k=n}^{m} \rho^k \circ \varepsilon_{t-ks}\Big] = \sum (\rho^{2k} \sigma_\varepsilon^2 + \rho^k (1-\rho^k) \mu_\varepsilon \leq \frac{\rho^{2n}}{1-\rho^2} \sigma_\varepsilon^2 + \frac{\rho^n}{1-\rho} \mu_\varepsilon$$

Since $\rho \in [0, 1)$, the right-hand sides in above equations tends to 0 as $n \to \infty$. This implies that the sequence $\{Z_{t,n}\}$ forms a Cauchy sequence in mean square sense and hence in probability. Therefore for all $t \in \mathbb{N}_0$, there is a random variable Z_t , which is the limit on the right-hand side in the equation, such that $Z_{t,n} \xrightarrow{p} Z_t$ as $n \to \infty$. Let the non-negative integer-valued stochastic process $\{X_t\}$ satisfy (2.2). By successively substituting we obtain

$$X_t = \rho \circ X_{t-s} + \varepsilon_t = \rho \circ (\rho \circ X_{t-2s} + \varepsilon_{t-s}) + \varepsilon_t$$

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

$$X_t \stackrel{d}{=} \rho^2 \circ X_{t-2s} + \rho \circ \varepsilon_{t-s} + \varepsilon_t = \rho^2 \circ X_{t-2s} + Z_{t,1},$$

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

$$X_t \stackrel{d}{=} \rho^n \circ X_{t-ns} + \sum_{k=0}^{n-1} \rho^k \circ \varepsilon_{t-ks} = \rho^n \circ X_{t-ns} + Z_{t,n-1},$$

If μ_x and σ_x^2 represent, respectively, the mean and variance of a stationary solution $\{X_t\}$ we obtain $\{X_t\}$

$$E[\rho^k \circ X_{t-ns}] = \rho^n \mu_x$$

$$Var[\rho^k \circ X_{t-ns}] = \rho^{2n}\sigma_x^2 + \rho^n(1 - \rho^n)\mu_x$$

Since $\rho \in [0, 1)$, we obtain

$$\lim_{n\to\infty} E[\rho^k \circ X_{t-ns}] = \lim_{n\to\infty} Var[\rho^k \circ X_{t-ns}] = 0$$

Therefore $\rho^n \circ X_{t-ns} \xrightarrow{P} 0$ as $n \to \infty$, and thus $Z_{t,n} \xrightarrow{d} X_t$ as $n \to \infty$ for all $t \in \mathbb{N}_0$, where $\xrightarrow{d}$ denotes convergence in distribution. Hence $X_t \stackrel{d}{=} Z_t$ for all $t \in \mathbb{N}_0$, which means the uniqueness of the stationary marginal solution. At last, it is showed that the distribution of the process $\{Z_t\}_{t\in\mathbb{N}_0}$ is the solution of the Equation (2.2). Using the properties of the binomial thinning operator, the following is derived:

$$Z_{t,n} = \sum_{k=0}^{n} \rho^k \circ \varepsilon_{t-ks} \stackrel{d}{=} \varepsilon_t + \sum_{k=0}^{n-1} \rho^{k+1} \circ \varepsilon_{t-(k+1)s} = \varepsilon_t + \rho \circ \left(\sum_{k=0}^{n-1} \rho^k \circ \varepsilon_{t-s-kz} \right) = \rho \circ Z_{t,n-1} + \varepsilon_t,$$

taking the limit in probability as $n \to \infty$ the equality $Z_t \stackrel{d}{=} \rho \circ Z_t + \varepsilon_t$ is obtained, that is the distribution of Z_t is a stationary marginal distribution to Equation (2.2).

References

[1] Aghababaei Jazi, M., Jones, G. and Lai, C.D., 2012. First-order integer valued AR processes with zero inflated Poisson innovations. *Journal of Time Series Analysis*, 33(6), pp.954-963. http://dx.doi.org/10.1111/j.1467-9892.2012.00809.x

[2] Aghababaei Jazi, M., Jones, G. and Lai, C.D., 2022. Integer valued AR (1) with geometric innovations. *Journal of the Iranian Statistical Society*, 11(2), pp.173-190.

[3] Al-Osh, M.A. and Alzaid, A.A., 1987. First-order integer-valued autoregressive (INAR (1)) process. *Journal of Time Series Analysis*, 8(3), pp.261-275. http://dx.doi.org/10.1111/j.1467-9574.1988.tb01521.x

[4] Barreto-Souza, W., 2015. Zero-modified geometric INAR (1) process for modelling count time series with deflation or inflation of zeros. *Journal of Time Series Analysis*, 36(6), pp.839-852. http://dx.doi.org/10.1111/jtsa.12131

[5] Barreto-Souza, W., 2019. Mixed poisson INAR (1) processes. *Statistical papers*, 60(6), pp.2119-2139. http://dx.doi.org/10.1007/s00362-017-0912-x

[6] Bourguignon, M., Rodrigues, J. and Santos-Neto, M., 2019. Extended Poisson INAR (1) processes with equidispersion, underdispersion and overdispersion. *Journal of Applied Statistics*, 46(1), pp.101-118. http://dx.doi.org/10.1080/02664763.2018.1458216

[7] Bourguignon, M., LP Vasconcellos, K., Reisen, V.A. and Ispány, M., 2016. A Poisson INAR (1) process with a seasonal structure. *Journal of Statistical Computation and Simulation*, 86(2), pp.373-387. http://dx.doi.org/10.1080/00949655.2015.1015127

[8] Bourguignon, M. and Vasconcellos, K.L., 2015. First order non-negative integer valued autoregressive processes with power series innovations. *Brazilian Journal of Probability and Statistics*, 29(1) , pp.71–93. http://dx.doi.org/10.1214/13-bjps229

[9] Bourguignon, M. and Weiß, C.H., 2017. An INAR (1) process for modeling count time series with equidispersion, underdispersion and overdispersion. *Test*, 26(4), pp.847-868. http://dx.doi.org/10.1007/s11749-017-0536-4

[10] Brannas, K., 1995. Explanatory variables in the AR (1) count data model. *Umea economic studies*, 381. http://dx.doi.org/10.2139/ssrn.1313842

[11] Buteikis, A. and Leipus, R., 2020. An integer-valued autoregressive process for seasonality. *Journal of Statistical Computation and Simulation*, 90(3), pp.391-411. http://dx.doi.org/10.1080/00949655.2019.1685995

[12] Delaporte, P., 1959. Quelques problèmes de statistique mathématique posés par l'assurance automobile et le bonus pour non sinistre. *Bulletin Trimestriel de l'Institut des Actuaires Français*, 227, pp.87-102.

[13] Ferland, R., Latour, A. and Oraichi, D., 2006. Integer-valued GARCH process. *Journal of time series analysis*, 27(6), pp.923-942. http://dx.doi.org/10.1111/j.1467-9892.2006.00496.x

[14] Fernández-Fontelo, A., Fontdecaba, S., Alba, A. and Puig, P., 2017. Integer-valued AR processes with Hermite innovations and time-varying parameters: An application to bovine fallen stock surveillance at a local scale. *Statistical modelling*, 17(3), pp.172-195. http://dx.doi.org/10.1177/1471082x16683113

[15] Freeland, R.K., 1998. *Statistical analysis of discrete time series with application to the analysis of workers' compensation claims data* (Doctoral dissertation, University of British Columbia). http://dx.doi.org/10.14288/1.0088709

[16] Freeland, R.K. and McCabe, B., 2005. Asymptotic properties of CLS estimators in the Poisson AR (1) model. *Statistics & probability letters*, 73(2), pp.147-153. http://dx.doi.org/10.1016/j.spl.2005.03.006

[17] Freeland, R.K. and McCabe, B.P., 2004. Analysis of low count time series data by Poisson autoregression. *Journal of time series analysis*, 25(5), pp.701-722. http://dx.doi.org/10.1111/j.1467-9892.2004.01885.x

[18] Jose, K.K. and Mariyamma, K.D., 2016. A note on an integer valued time series model with Poisson–negative binomial marginal distribution. *Communications in Statistics-Theory and Methods*, 45(1), pp.123-131. http://dx.doi.org/10.1080/03610926.2013.826979

[19] Jung, R.C., Ronning, G. and Tremayne, A.R., 2005. Estimation in conditional first order autoregression with discrete support. *Statistical papers*, 46, pp.195-224. http://dx.doi.org/10.1007/bf02762968

[20] Kim, H.Y. and Park, Y., 2008. A non-stationary integer-valued autoregressive model. *Statistical papers*, 49, pp.485-502. http://dx.doi.org/10.1007/s00362-006-0028-1

[21] Kim, H. and Lee, S., 2017. On first-order integer-valued autoregressive process with Katz family innovations. *Journal of Statistical Computation and Simulation*, 87(3), pp.546-562. http://dx.doi.org/10.1080/00949655.2016.1219356

[22] Klimko, L.A. and Nelson, P.I., 1978. On conditional least squares estimation for stochastic processes. *The Annals of statistics*, pp.629-642. http://dx.doi.org/10.1214/aos/1176344207

[23] McKenzie, E., 1985. Some simple models for discrete variate time series 1. *JAWRA Journal of the American Water Resources Association*, 21(4), pp.645-650. http://dx.doi.org/10.1111/j.1752-1688.1985.tb05379.x

[24] Quoreshi, S., Uddin, R. and Mamode Khan, N., 2020. A review of INMA integer-valued model class, application and further development. *Filomat*, 34(1), pp.143-152. http://dx.doi.org/10.2298/fil2001143q

[25] Scotto, M.G., Weiss, C.H. and Gouveia, S., 2015. Thinning-based models in the analysis of integer-valued time series: a review. *Statistical Modelling*, 15(6), pp.590-618. http://dx.doi.org/10.1177/1471082x15584701

[26] Shalbaf, M., Parham, G. and Chinipardaz, R., 2022. Binomial Thinning Integer-Valued AR (1) with Poisson– Fold Zero Modified Geometric Innovations. *Journal of Sciences, Islamic Republic of Iran*, 33(1), pp.55-63. https://doi.org/10.22059/jsciences.2021.320996.1007633

[27] Steutel, F.W. and van Harn, K., 1979. Discrete analogues of self-decomposability and stability. *The Annals of Probability*, pp.893-899. http://dx.doi.org/10.1214/aop/1176994950

[28] Silva, M.E. and Oliveira, V.L., 2005. Difference equations for the higher order moments and cumulants of the INAR (p) model. *Journal of Time Series Analysis*, 26(1), pp.17-36. http://dx.doi.org/10.1111/j.1467-9892.2005.00388.x

[29] Tian, S., Wang, D. and Cui, S., 2020. A seasonal geometric INAR process based on negative binomial thinning operator. *Statistical Papers*, 61, pp.2561-2581. http://dx.doi.org/10.1007/s00362-018-1060-7

[30] Weiß, C.H., 2008. Thinning operations for modeling time series of counts—a survey. *AStA Advances in Statistical Analysis*, 92, pp.319-341. http://dx.doi.org/10.1007/s10182-008-0072-3

[31] Weiß, C.H., 2009. Controlling jumps in correlated processes of Poisson counts. *Applied Stochastic Models in Business and Industry*, 25(5), pp.551-564. https://doi.org/10.1002/asmb.744

[32] Weiß, C.H., 2009. Modelling time series of counts with overdispersion. *Statistical Methods and Applications*, 18, pp.507-519. https://doi.org/10.1007/s10260-008-0108-6

[33] Willmot, G.E. and Sundt, B., 1989. On evaluation of the Delaporte distribution and related distributions. *Scandinavian Actuarial Journal*, 1989(2), pp.101-113. https://doi.org/10.1080/03461238.1989.10413859

[34] Zheng, H., Basawa, I.V. and Datta, S., 2006. Inference for pth-order random coefficient integer-valued autoregressive processes. *Journal of Time Series Analysis*, 27(3), pp.411-440. https://doi.org/10.1111/j.1467-9892.2006.00472.x

[35] Zhu, R. and Joe, H., 2006. Modelling count data time series with Markov processes based on binomial thinning. *Journal of Time Series Analysis*, 27(5), pp.725-738. https://doi.org/10.1111/j.1467-9892.2006.00485.x

Maryam Shalbaf
Department of Statistics,
Faculty of Mathematical Sciences and Computer,
Shahid Chamran University of Ahvaz,
Ahvaz, Iran
Email: Shalbaf.maryam@gmail.com

Gholamali Parham
Department of Statistics,
Faculty of Mathematical Sciences and Computer,
Shahid Chamran University of Ahvaz,
Ahvaz, Iran
Email: Parham_g@scu.ac.ir

Rahim Chinipardaz
Department of Statistics,
Faculty of Mathematical Sciences and Computer,
Shahid Chamran University of Ahvaz,
Ahvaz, Iran
Email: Chinipardaz_r@scu.ac.ir