**Journal of Sciences, Islamic Republic of Iran 33(1): 55 - 63 (2022)** http://jsciences.ut.ac.ir
University of Tehran, ISSN 1016-1104

# Binomial Thinning Integer-Valued AR (1) with Poisson – α Fold Zero Modified Geometric Innovations

M. Shalbaf, G.A. Parham*, and R. Chinipardaz

*Department of Statistics, Faculty of Mathematical Sciences and Computer Shahid Chamran University of Ahvaz, Ahvaz, Islamic Republic of Iran*

Received: 23 Mar 2021 / Revised: 10 Oct 2021 / Accepted: 21 Dec 2021

## Abstract

Real count data time series often show the phenomenon of the overdispersion. In this paper, we introduce the first-order integer-valued autoregressive process. The univariate *marginal distribution* is derived from the Delaporte distribution and the innovations are convolution of Poisson with $\alpha$-fold zero modified geometric distribution, based on binomial thinning operator, for modelling integer-valued time series with overdispersion. Some properties of the model are derived. The methods of Yule–Walker, conditional lea st squares and conditional maximum likelihood are used for estimating of the parameters, and their asymptotic properties are established. The Monte Carlo experiment is conducted to evaluate the performances of these estimators in finite samples. The model is fitted to time series of the weekly number of syphilis cases that are overdispersed count data.

**Keywords:** $\alpha$-fold zero modified geometric; Binomial thinning; Count time series; Delaporte distribution; INAR (1) models.

## Introduction

During the last decades, modelling count time series has been considered in many articles often as counts of events or individuals in consecutive time intervals. For example, they arise as the number of births at a hospital in successive months, the number of road accidents and the number of diseases in a certain area in successive months. The integer-valued autoregressive (INAR) time series models are constructed usually based on the binomial thinning operator that introduced by Steutel and van Harn [1], which is defined as follows:

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

and 0 otherewise, where the counting series $Y \coloneqq \{Y_i\}_{i\geq 1}$ is a sequence of independent identically distributed Bernoulli random variables with fixed success probability $\rho \in [0,1]$ and $X$ is a non-negative integer valued random variable independent of $Y$.

McKenzie [2] and Al-Osh and Alzaid [3] introduced the INAR(1) model as follows

$$X_t = \rho \circ X_{t-1} + \varepsilon_t, \qquad t \in Z, \tag{2}$$

where $0 \leq \rho < 1$, $\{\varepsilon_t\}_{t\in Z}$ is a sequence of independent and identically distributed integer valued random variables, called innovations and for each $t$, $\varepsilon_t$ is independent of $X_{t-l}$ for all $l \geq 1$, $E(\varepsilon_t) = \mu_\varepsilon$ and $Var(\varepsilon_t) = \sigma_\varepsilon^2$. From the results of Al-Osh and Alzaid [3], we have $\rho \in [0,1)$ and $\rho = 1$ are the conditions of (strictly) stationarity and non-stationarity of the process $\{X_t\}_{t\in Z}$, respectively. The autocorrelation function

* Corresponding author: Tel: +986133331041; Email: Parham_g@scu.ac.ir

(ACF) of model (2) is $\rho_k = Corr(X_t, X_{t-k}) = \rho^k$, for $k \geq 0$, that is, it is of AR(1) type, but the only non-negative autocorrelation is allowed. Also, $\rho = 0$ ($\rho > 0$) implies the independence (dependence) of the observations of $\{X_t\}_{t\in Z}$. The mean and variance of $\{X_t\}_{t\in Z}$ are given by

$\mu_X := E(X_t) = \frac{\mu_\varepsilon}{1-\rho}$ and $\sigma_X^2 := Var(X_t) = \frac{\rho\mu_\varepsilon+\sigma_\varepsilon^2}{1-\rho^2}$.

A commonly used variability measure of a random variable is the Fisher index of dispersion, defined by $FI_X = Var(X)/E(X)$, that is a measure of aggregation or disaggregation, for more details see [4], page 163. Thus, the Fisher index of dispersion of $\{X_t\}_{t\in Z}$ in Equation (2) is given by

$$FI_X = \frac{FI_\varepsilon+\rho}{1+\rho}. \quad (3)$$

If $\varphi_X(s)$ and $\varphi_\varepsilon(s)$ denote the probability generating function (pgf) of $\{X_t\}_{t\in Z}$ and $\{\varepsilon_t\}_{t\in Z}$, respectively, then the stationary marginal distribution of $\{X_t\}_{t\in Z}$ can be determined from the equation $\varphi_X(s) = \varphi_X(1-\rho+\rho s).\varphi_\varepsilon(s)$, This allows for various types of marginal distributions, including the Poisson [3], geometric [5], generalized Poisson [6] and Poisson-geometric distributions [7] .

Overdispersion is an important concept in the analysis of discrete data. The Poisson INAR process is not suitable for modelling overdispersed counts because Poisson distribution is equidispersed. Various methods have been proposed to overcome it. A simple approach is to change the innovation distribution. Another suggestion to deal with overdispersion is to change the type of thinning operator, which often makes change the distribution of innovation. The third approach is to change the marginal distribution of the process. A reason for overdispersion, reported in the literature, is the presence of a positive correlation between the monitored events [8-9]. Jazi, Jones and Lai [10] introduced the INAR(1) process with geometric innovations. Jazi, Jones and Lai [11] discussed an INAR(1) process with zero-inflated Poisson innovations. Barreto-Souza [12] proposed INAR processes with zero-modified geometric marginals. Bourguignon and Vasconcellos [13] studied a new stationary INAR(1) process with power series innovations. Jose, and Mariyamma [14] introduced an INAR(1) model with Poisson-negative binomial marginal distribution. Fernáandez-Fontelo Fontdecaba, Alba, Puig [15] introduced a generalization of the classical Poisson-based INAR models whose innovations follow a Hermite distribution. Kim and Lee [16] considered the INAR (1) process with Katz family innovations. An INAR (1) process for modelling count time series with equidispersion, underdispersion and overdispersion was studied by Bourguignon and Weiß [17]. Bourguignon, Rodrigues and Santos-Neto [18] introduced two new binomial thinning INAR (1) processes with double Poisson (DP) and GP innovations, denoted by INARDP (1) and INARGP (1), respectively, for modelling non-negative integer-valued time series with equidispersion, underdispersion or overdispersion. A common way for treating overdispersion in count data is to use the mixed Poisson distributions, which is obtained by introducing a latent random effect on the mean of a Poisson distribution. Barreto-Souza [19] proposed class of overdispersed INAR (1) processes with marginals belonging to a general class of mixed Poisson distributions.

The Delaporte distribution is a discrete probability distribution that has received attention in actuarial science. It can be defined using the convolution of a negative binomial distribution with a Poisson distribution. Just as the negative binomial distribution can be viewed as a Poisson distribution where the mean parameter is itself a random variable with a gamma distribution, the Delaporte distribution can be viewed as a compound distribution based on a Poisson distribution, where there are two components to the mean parameter: a fixed component, which has the λ parameter, and a gamma-distributed variable component, which has the α and β parameters. The distribution is named for Pierre Delaporte, who analyzed it in relation to automobile accident claim counts in 1959. For the special value of parameters this distribution reduces to Poisson, Polya, or geometric distribution.

In this paper, we estimate the unknown parameters of a first-order integer-valued autoregressive that the univariate *marginal distribution* is derived from the Delaporte distribution and the innovations are convolution of Poisson with $\alpha$-fold zero modified geometric distribution. We denote this model by DELINAR (1). This model is suitable for modelling non-negative integer-valued time series with overdispersion. The article is organized as follows. The model is defined in Section 2 and we derive the transition probabilities of the model based on marginal and innovation mass function. Also, some of properties model are outlined. In Section 3, estimation methods for the model parameters and asymptotic distribution for some parameters are discussed. Section 4 discusses some simulation results for the estimation methods. In Section 5, the model is applied to a well-known data set. Finally, we conclude in Section 6.

***INAR (1) process and its marginals***

In this section, we study structural properties of this 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.

***The Delaporte distribution***

The Delaporte distribution is a Poisson mixture proposed to fit the number of claims in an insurance portfolio [20]. The Delaporte distribution has probability generating function

$$G_X(t) = e^{-\lambda(1-t)}\left(\frac{1}{1+\beta(1-t)}\right)^{\alpha}, \qquad (4)$$

where $|t| \leq 1$, $\lambda > 0$, $\alpha, \beta > 0$. This shows that it is the convolution of a negative binomial and Poisson random variables (see, [20] and [21]). The probability mass function (pmf) corresponding to (4) 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 $\mathrm{Del}(\lambda, \alpha, \beta)$.

Differentiating the probability generating function of Delaporte distribution, it is easy to show

$$\mu_X := E(X) = \lambda + \alpha\beta \qquad (5)$$

and

$$\sigma_X^2 := Var(X) = \lambda + \alpha\beta(1+\beta). \qquad (6)$$

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 this distribution shows overdispersion.

***INAR(1) process with convolution Poisson with α-fold zero modified geometric distribution***

In this section we consider a stationary integer valued process $\{X_t\}_{t\in Z}$ in (2) with $\mathrm{Del}(\lambda, \alpha, \beta)$ marginal distribution, where $\rho \in (0, 1)$, and $\{\varepsilon_t\}_{t\in Z}$ is a sequence of i.i.d. random variables. Let $\Phi_X(s)$ and $\Phi_\varepsilon(s)$ be the alternative probability generating function of the random variables $X_t$ and $\varepsilon_t$, respectively. From (2) and the stationarity of the process $\{X_t\}_{t\in Z}$ it follows that the random variable $\varepsilon_t$ has the alternative probability generating function (apgf)

$$\Phi_{\varepsilon_t}(s) = \frac{\Phi_{X_t}(s)}{\Phi_{X_{t-1}}(\rho s)} = e^{-\lambda(1-\rho)s}\left[\rho + (1-\rho)\frac{1}{1+\beta s}\right]^{\alpha}.$$

Thus, it follows that the innovations sequence $\{\varepsilon_t\}_{t\in Z}$ can be represented as $\varepsilon_t = Y_1 + Y_2$, where $Y_1 \sim Po(\lambda(1-\rho))$ and $Y_2$ are $\alpha$-fold convolutions of zero-modified geometric distribution. Therefore, the expectation and the variance of the random variable $\varepsilon_t$ are

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

and

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

The conditional distribution of $X_t$ given $X_{t-1}$ has the form

$$p_{ij} = P(X_t = j | X_{t-1} = i) = P(\rho \circ X_{t-1} + \varepsilon_t = x_t | X_{t-1} = x_t).$$

Now, $B_{X_{t-1}}^{\rho} := \rho \circ X_{t-1} | X_{t-1} \sim \mathrm{binomial}(X_{t-1}, \rho)$, also $\varepsilon_t$ is independent of $X_{t-1}$ and is convolution of two random variables with distribution of $P_{(\lambda,\rho)} \sim Po(\lambda(1-\rho))$ and $\alpha$-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

$$P(X_t = x_t | X_{t-1} = x_{t-1}) = P\left(B_{X_{t-1}}^{\rho} + P_{(\lambda,\rho)} + \alpha FZMG_{(p,p_0)} = x_t\right)$$
$$= \sum_{s=0}^{x_t} P(\alpha FZMG = s) \sum_{d=0}^{\min\{x_t-s, x_{t-1}\}} P(B_{X_{t-1}}^{\rho} = d) P\left(P_{(\lambda,\rho)} = x_t - s - d\right), \quad (7)$$

where,

$$P\left(P_{(\lambda,\rho)} = u\right) = \frac{e^{-\lambda(1-\rho)}[\lambda(1-\rho)]^u}{u!}, \quad u = 0,1,2,\cdots,$$

$$P\left(B_{X_{t-1}}^{\rho} = i\right) = \binom{x_{t-1}}{i}\rho^i(1-\rho)^{x_{t-1}-i}, \quad i = 0,1,\cdots,x_{t-1},$$

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

$$P(Z_i = k) = \begin{cases} P_0 & k = 0 \\ (1-\rho)\dfrac{\beta^k}{(1+\beta)^{k+1}} & k = 1,2,\cdots \end{cases}$$

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

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

$E(X_t|X_{t-1}) = \rho X_{t-1} + (1-\rho)\mu_X$

and

$Var(X_t|X_{t-1}) = \rho(1-\rho)X_{t-1} + (1-\rho^2)\sigma_X^2 - \rho(1-\rho)\mu_X$.

***Parameter estimation and asymptotic properties***

Assume that we have n observations $X_1, X_2, \dots, X_n$ from DELINAR(1) process. In the DELINAR(1) model we have four parameters, we assume that $\alpha$ is known, therefore, three parameters $\rho$, $\lambda$ and $\beta$ 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).

***Yule-Walker estimation***

Let us consider the Yule–Walker estimators of the unknown parameters $\rho$, $\lambda$ and $\beta$. Because $\rho = \rho_1$, $E(X_t) = \lambda + \alpha\beta$ and $Var(X_t) = \lambda + \alpha\beta(1+\beta)$, we have that the Yule–Walker estimators are

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

$$\hat{\lambda}_{YW} = \bar{X} - \alpha\sqrt{\frac{S^2-\bar{X}}{\alpha}},$$

$$\hat{\beta}_{YW} = \sqrt{\frac{S^2-\bar{X}}{\alpha}},$$

where $\bar{X}$ and $S^2$ are the sample mean and sample variance, respectively.

***Conditional least squares estimation***

We derive the conditional least squares estimators of the parameters $\rho$, $\mu$ and $\sigma^2$, where $\mu = E(X_t) = \lambda + \alpha\beta$ and $\sigma^2 = Var(X_t) = \lambda + \alpha\beta(1+\beta)$. We use the two-steps conditional least squares method [22]. In the first step, we derive the conditional least squares estimators of the parameters $\rho$ and $\mu$. They are obtained by minimizing the sum of squares

$$S_{1CLS}(\rho,\mu) = \sum_{t=2}^{n} e_{1t}^2 = \sum_{t=2}^{n}(X_t - E[X_t|X_{t-1}])^2 = \sum_{t=2}^{n}(X_t - \rho X_{t-1} - (1-\rho)\mu)^2.$$

Hence, the conditional least squares estimators of the parameters $\rho$ and $\mu$ are

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

and

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

In the second step, we consider the estimation of the parameter $\sigma^2$. We define a new random variable $V_t$ as $V_t = (X_t - E(X_t|X_{t-1}))^2 = (X_t - \rho X_{t-1} - (1-\rho)\mu)^2$. It is easy to show that $E(V_t|X_{t-1}) = Var(X_t|X_{t-1}) = \rho(1-\rho)X_{t-1} + (1-\rho^2)\sigma^2 - \rho(1-\rho)\mu$.

Now, the conditional least squares estimator of the parameter $\sigma^2$ can be obtained by minimizing the sum of squares

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

Thus, the conditional least squares estimator of the parameter $\sigma^2$ is

$$\hat{\sigma}_{cls}^2 = \frac{\sum_{t=2}^{n}(X_t - \hat{\rho}_{cls}X_{t-1} - (1-\hat{\rho}_{cls})\hat{\mu}_{cls})^2 - \hat{\rho}_{cls}(1-\hat{\rho}_{cls})\sum_{t=2}^{n}(X_{t-1}-\hat{\mu}_{cls})}{(n-1)(1-\hat{\rho}_{CLS}^2)}.$$

Finally, the conditional least squares estimators for the parameters $\lambda$, $\beta$ are, respectively,

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

**Proposition.** The estimators $\hat{\rho}_{cls}$ and $\hat{\mu}_{cls}$ are strongly consistent for $\rho$ and $\mu$ respectively and satisfy the asymptotic normality

$$\sqrt{n}[(\hat{\rho}_{cls}, \hat{\mu}_{cls})^T - (\rho,\mu)^T] \xrightarrow{d} N_2\left((0,0)^T, V^{-1}WV^{-1}\right)$$

as n→ ∞, where

$$V^{-1} = \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{(1-\rho)^2} \end{pmatrix}$$

and W being a 2×2 symmetric matrix given by

$$W = (1-\rho)\begin{pmatrix} \rho\mu_*^3 & \rho(1-\rho)\sigma^2 \\ \rho(1-\rho)\sigma^2 & (1-\rho)(1-\rho^2)\sigma^2 \end{pmatrix}$$

where $\mu_*^3$ is third central moment of $X_t$.

**Proof.** It can be verified that the regularity conditions given in Theorem 3.2 of [23], are satisfied by DELINAR(1) process. So, the proof is achieved.

***Conditional maximum likelihood estimation***

Let $X_1, X_2, \dots, X_n$ be a random sample of size n from a stationary DELINAR(1) process with parameters $\rho$, $\lambda$ and $\beta$. The conditional log-likelihood function is given by

$$CL(\rho,\lambda,\beta) = \sum_{t=2}^{n} \log P(X_t = x_t|X_{t-1} = x_{t-1}),$$

where $P(X_t = j|X_{t-1} = i)$ is defined in (7). The conditional maximum likelihood estimators are obtained by maximizing $CL(\rho, \lambda, \beta)$. In practice, there will be no closed form for the CML estimates and numerical methods are needed.

### *Simulation*

In this section we show some simulation results for the different values of the parameters. All simulations were carried out using the R programming language. We generate the sample sizes n = 100, 200, 400, 800 from the DELINAR(1) process, and the number of Monte Carlo replications for each case is 1000. In this simulation, we set (a) $(\rho, \beta, \lambda) = (0.1,2,6)$, (b) $(\rho, \beta, \lambda) = (0.2,3,1)$, (c) $(\rho, \beta, \lambda) = (0.5,1,3)$ and (d) $(\rho, \beta, \lambda) = (0.8,1.5,1)$. Tables 1, 2, 3 and 4 show the empirical bias and mean square error (MSE) of the

**Table 1.** Empirical bias and MSE (in parentheses) of estimators of parameters for $(\rho, \beta, \lambda) = (0.1,2,6)$.

| | Estimator of ρ | | | Estimator of β | | | Estimator of λ | | |
|---|---|---|---|---|---|---|---|---|---|
| **n** | $\hat{\rho}_{YW}$ | $\hat{\rho}_{CLS}$ | $\hat{\rho}_{CML}$ | $\hat{\beta}_{YW}$ | $\hat{\beta}_{CLS}$ | $\hat{\beta}_{CML}$ | $\hat{\lambda}_{YW}$ | $\hat{\lambda}_{CLS}$ | $\hat{\lambda}_{CML}$ |
| **100** | -0.015 | -0.012 | -0.001 | -0.069 | -0.093 | -0.110 | 0.099 | 0.147 | 0.181 |
| | (0.010) | (0.010) | (0.006) | (0.173) | (0.180) | (0.250) | (0.626) | (0.650) | (0.881) |
| **200** | -0.012 | 0.011 | -0.008 | -0.002 | -0.013 | -0.005 | 0.009 | 0.034 | 0.018 |
| | (0.005) | (0.005) | (0.004) | (0.077) | (0.077) | (0.073) | (0.298) | (0.300) | (0.286) |
| **400** | -0.007 | -0.007 | -0.006 | -0.014 | -0.019 | -0.014 | 0.024 | 0.035 | 0.024 |
| | (0.002) | (0.002) | (0.002) | (0.036) | (0.037) | (0.036) | (0.145) | (0.145) | (0.143) |
| **800** | -0.002 | -0.002 | -0.003 | -0.005 | -0.007 | -0.006 | 0.004 | 0.010 | 0.007 |
| | (0.001) | (0.001) | (0.001) | (0.018) | (0.018) | (0.016) | (0.065) | (0.066) | (0.061) |

**Table 2.** Empirical bias and MSE (in parentheses) of estimators of parameters for $(\boldsymbol{\rho}, \boldsymbol{\beta}, \boldsymbol{\lambda}) = (\mathbf{0.2}, \mathbf{3}, \mathbf{1})$.

| | Estimator of ρ | | | Estimator of β | | | Estimator of λ | | |
|---|---|---|---|---|---|---|---|---|---|
| **n** | $\hat{\rho}_{YW}$ | $\hat{\rho}_{CLS}$ | $\hat{\rho}_{CML}$ | $\hat{\beta}_{YW}$ | $\hat{\beta}_{CLS}$ | $\hat{\beta}_{CML}$ | $\hat{\lambda}_{YW}$ | $\hat{\lambda}_{CLS}$ | $\hat{\lambda}_{CML}$ |
| **100** | -0.022 | -0.018 | -0.008 | -0.057 | -0.079 | -0.048 | 0.113 | 0.156 | 0.096 |
| | (0.009) | (0.009) | (0.005) | (0.093) | (0.198) | (0.146) | (0.552) | (0.567) | (0.329) |
| **200** | -0.006 | -0.004 | -0.001 | -0.033 | -0.043 | -0.028 | 0.057 | 0.077 | 0.047 |
| | (0.004) | (0.004) | (0.002) | (0.096) | (0.097) | (0.068) | (0.284) | (0.286) | (0.155) |
| **400** | -0.005 | -0.004 | -0.002 | -0.007 | -0.012 | -0.009 | 0.013 | 0.025 | 0.018 |
| | (0.002) | (0.002) | (0.001) | (0.051) | (0.051) | (0.036) | (0.139) | (0.140) | (0.074) |
| **800** | -0.002 | -0.002 | -0.000 | -0.007 | -0.009 | -0.005 | 0.016 | 0.020 | 0.013 |
| | (0.001) | (0.001) | (0.000) | (0.026) | (0.026) | (0.017) | (0.074) | (0.074) | (0.037) |

**Table 3.** Empirical bias and MSE (in parentheses) of estimators of parameters for $(\rho, \beta, \lambda) = (0.5,1,3)$.

| | Estimator of **ρ** | | | Estimator of **β** | | | Estimator of **λ** | | |
|---|---|---|---|---|---|---|---|---|---|
| **n** | $\hat{\rho}_{YW}$ | $\hat{\rho}_{CLS}$ | $\hat{\rho}_{CML}$ | $\hat{\beta}_{YW}$ | $\hat{\beta}_{CLS}$ | $\hat{\beta}_{CML}$ | $\hat{\lambda}_{YW}$ | $\hat{\lambda}_{CLS}$ | $\hat{\lambda}_{CML}$ |
| **100** | -0.023 | -0.014 | -0.008 | -0.042 | -0.062 | -0.036 | 0.091 | 0.132 | 0.078 |
| | (0.008) | (0.007) | (0.005) | (0.107) | (0.115) | (0.101) | (0.496) | (0.532) | (0.463) |
| **200** | -0.013 | -0.009 | -0.006 | -0.045 | -0.057 | -0.042 | 0.108 | 0.130 | 0.100 |
| | (0.004) | (0.003) | (0.002) | (0.067) | (0.069) | (0.056) | (0.280) | (0.293) | (0.236) |
| **400** | -0.005 | -0.003 | -0.001 | -0.014 | -0.019 | -0.009 | 0.037 | 0.045 | 0.026 |
| | (0.001) | (0.001) | (0.001) | (0.033) | (0.034) | (0.024) | (0.143) | (0.145) | (0.106) |
| **800** | -0.004 | -0.003 | -0.000 | -0.014 | -0.016 | -0.007 | 0.025 | 0.030 | 0.011 |
| | (0.001) | (0.001) | (0.000) | (0.017) | (0.017) | (0.012) | (0.069) | (0.070) | (0.053) |

**Table 4.** Empirical bias and MSE (in parentheses) of estimators of parameters for $(\rho, \beta, \lambda) = (0.8,1.5,1)$.

| | **Estimator of ρ** | | | **Estimator of β** | | | **Estimator of λ** | | |
|---|---|---|---|---|---|---|---|---|---|
| **n** | $\hat{\rho}_{YW}$ | $\hat{\rho}_{CLS}$ | $\hat{\rho}_{CML}$ | $\hat{\beta}_{YW}$ | $\hat{\beta}_{CLS}$ | $\hat{\beta}_{CML}$ | $\hat{\lambda}_{YW}$ | $\hat{\lambda}_{CLS}$ | $\hat{\lambda}_{CML}$ |
| **100** | -0.046 | -0.028 | -0.003 | -0.152 | -0.165 | 0.000 | 0.339 | 0.375 | 0.037 |
| | (0.006) | (0.004) | (0.000) | (0.274) | (0.284) | (0.161) | (0.959) | (1.037) | (0.307) |
| **200** | -0.028 | -0.020 | -0.002 | -0.124 | -0.132 | -0.014 | 0.262 | 0.283 | 0.042 |
| | (0.003) | (0.002) | (0.000) | (0.169) | (0.173) | (0.075) | (0.611) | (0.620) | (0.145) |
| **400** | -0.012 | -0.008 | -0.000 | -0.041 | -0.044 | 0.001 | 0.099 | 0.106 | 0.013 |
| | (0.001) | (0.001) | (0.000) | (0.085) | (0.086) | (0.038) | (0.283) | (0.287) | (0.070) |
| **800** | -0.005 | -0.003 | -0.000 | -0.025 | -0.027 | -0.005 | 0.058 | 0.060 | 0.015 |
| | (0.000) | (0.000) | (0.000) | (0.039) | (0.039) | (0.017) | (0.136) | (0.137) | (0.032) |

estimators obtained from the YW, CLS and CML methods. These tables show that the bias and standard error of the estimates of the parameters decrease as the sample size increases for all cases. As can be seen from the tables, CLS and YW methods show similar MSE behaviors. However the CML estimators have the best implementation on empirical bias and MSE compared with the YW and CLS estimators, because both biases and MSE for the CML estimators are smaller than those for the other methods. The bias of the estimators of $\rho$ and $\beta$ are negative, so they tend to underestimate of the parameters, and the bias of the estimators of $\beta$ is positive, so it tends to overestimate the parameter.

***Real data***

The application of the model was illustrated in this section by a real data set. This data set gives the number of syphilis cases in the United States monthly from 2007–2010 in Mid-Atlantic state given in tsinteger package [24] available for download at data (syphillis). The data consist of 209 observations, and they were already analyzed in [25] and [18].

The sample mean is 24.63, the sample variance is 105.68, and the first-order autocorrelation is 0.2322. The empirical Fisher index of dispersion is 4.29. The sample variance is much larger than the sample mean. Hence, the data seems to be overdispersed. Figure 1 shows the time series plot, the ACF and partial ACF (PACF). The ACF plot indicates that an integer-valued AR(1) my be suitable.

Table 5 gives the CML estimates (with corresponding standard errors in parentheses), Akaike information criterion (AIC) and Bayesian information criterion (BIC) for the fitted models. The values of the AIC and BIC are smaller for the DELINAR(1) model compared to those values of the INARP(1) model. Then, the suitable model by CML estimation is

$$X_t = 0.27 \circ X_{t-1} + \varepsilon_t, \qquad t \in Z,$$

where $X_t$ follows a Del(5.63, 2, 6.07) and $\varepsilon_t$ is a convolution of two random variables with distribution of Poisson(4.11) and zero-modified geometric with parameters $p = 0.37$, and $p_0 = 0.14$. The observed, fitted and forecasting values are shown in Figure 2 with

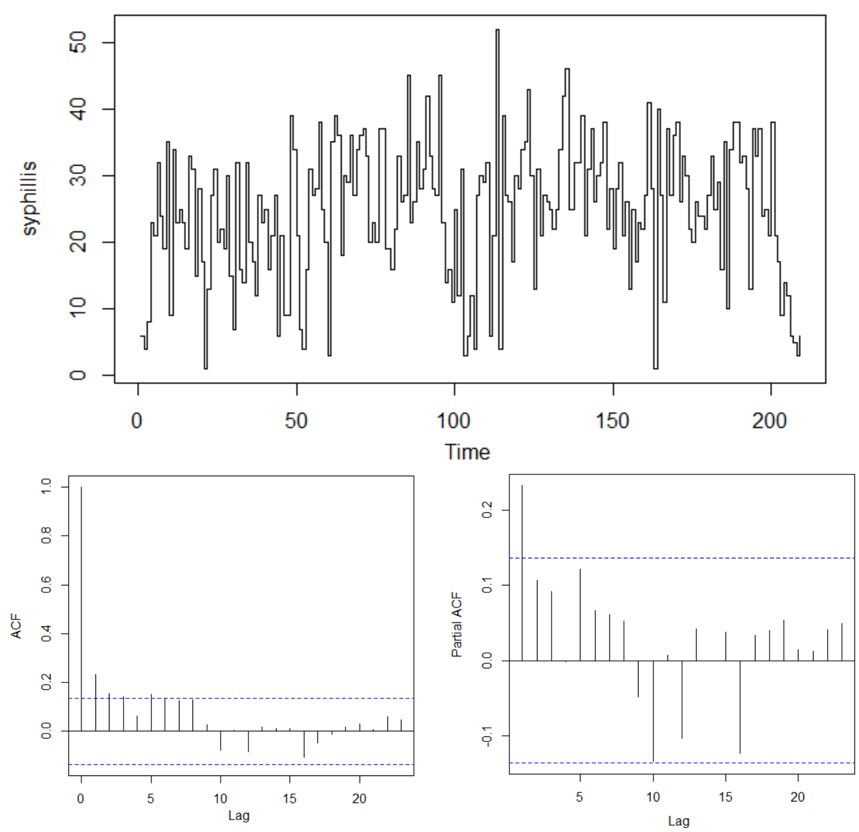


**Figure 1.** Time series of the number of syphilis cases and its sample ACF and PACF

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

| Model | Parameter | CML estimate | AIC | BIC | $\mu_x$ | $\sigma_x^2$ | $FI_x$ |
|---|---|---|---|---|---|---|---|
| **DELINAR (1)** | $\rho$ | 0.2704 (0.0008) | 1642.159 | 1652.186 | 17.7632 | 91.3598 | 5.143 |
| | $\beta$ | 6.0660 (0.0033) | | | | | |
| | $\lambda$ | 5.6312 (0.0039) | | | | | |
| **PINAR (1)** | $\rho$ | 0.1480 (0.0261) | 2016.54 | 2023.22 | 24.72 | 24.72 | 1 |
| | $\lambda$ | 21.063 (0.7087) | | | | | |

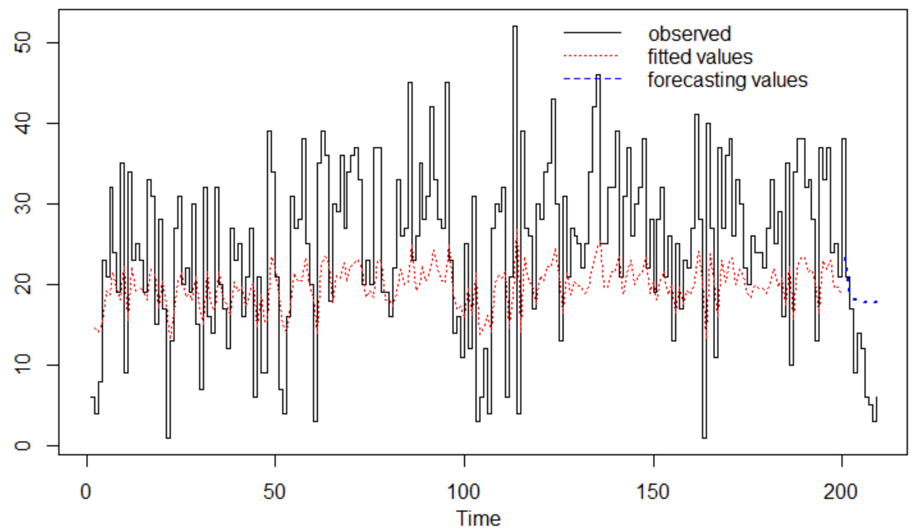


**Figure 2.** The plot of observed, fitted and forecasting values.

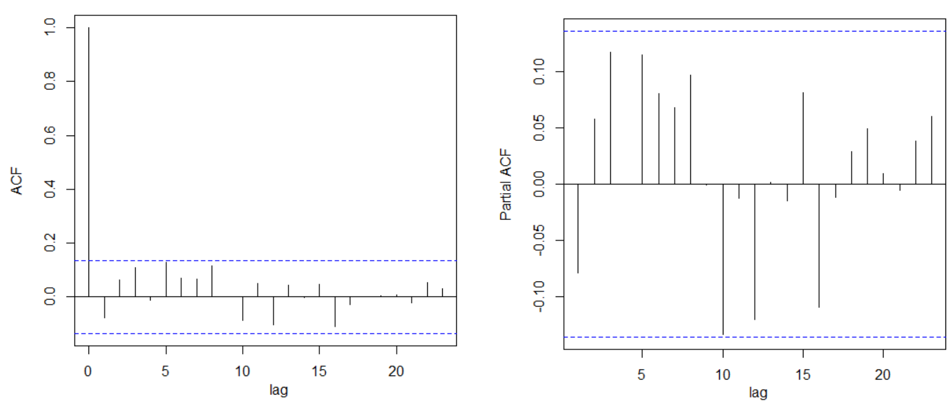


**Figure 3.** The ACF and PACF of residuals.

black, red and blue lines, respectively. Figure 3 shows ACF and PACF of residuals, and there are not serial correlations in the residuals. It means that the proposed model was suitable to fit the data.

## Results and Discussion

We discussed a stationary first-order nonnegative integer-valued autoregressive model for count data process based on binomial thinning operator. In this model the distribution of the innovations are convolution of a Poisson and $\alpha$-fold zero modified geometric distribution or marginal distribution of $X_t$ is Delaporte distribution This model is suitable for modelling overdispersed count time series data.

Afterward, we obtained some of the properties of the model and calculate YW, CLS and CML estimators of the parameters. The conditional maximum likelihood estimators of parameters have not closed form, therefore, numerical methods are needed. By using criterion function of the CLS, we see that the parameters of $\lambda$ and $\beta$ are not estimable, so we use the reparametrization, $\mu = \lambda + \alpha\beta$ and $\sigma^2 = \lambda + \alpha\beta(1+\beta)$. In the first step, by the conditional mean prediction error estimate parameter of the model, $\rho$, and the mean of $X_t$. To obtain an estimate of $\sigma^2$ in the second step, we use the normal equations based on the conditional variance prediction error. Since there is relation between parameters, we can find only joint asymptotic distribution of $\rho$ and $\mu$ that estimate based on criterion function of the CLS. In the simulation study, we compare YW, CLS and CML estimators. The simulation results show that the YW and CLS methods produce estimators with similar performances and that CML is better. Thus, we recommend the use of the CML method to estimate the model parameters of the DELINAR(1) process. Finally, we fitted the model to real data set to show the number of syphilis cases in the United States monthly from 2007–2010. The result fitted model shows that based on the AIC, BIC criterion, our model is better, compared to those values of the INARP(1) model. Also, observed, fitted and forecasted values show in the Figure 3. In this article, we have assumed $\alpha = 2$. It can be increased, or it can be considered unknown. It would be interesting to extend the model to the autoregressive model of order greater than 1.

***Funding***

We are grateful to the Research Council of Shahid Chamran University of Ahvaz for financial support (grant number:SCU.MS1400.800).

## Acknowledgement

The authors are grateful to editor and the reviewers for their valuable encouraging comments and suggestions.

## References

1. Steutel FW, Van Harn K. Discrete analogues of self-decomposability and stability. Ann Probab. 1979; 7(5): 893–899.
2. McKenzie E. Some simple models for discrete variate time series. J Am Water Resour Assoc. 1985; 21(4): 645–650.
3. Al-Osh MA, Alzaid AA. First-order interger-valued autoregressive (INAR(1)) process. J Time Ser Anal. 1987; 8(3): 261–275.
4. Johnson NL, Kemp AW, Kotz S. Univariate Discrete distributions. 3th ed. New York: John Wiley & Sons; 2005.
5. Al-Osh MA, Alzaid AA. Integer-valued moving average (INMA) process. Stat Pap. 1988; 29: 281–300.
6. Weiß CH. Thinning operations for modeling time series of counts – a survey. Adv Stat Anal. 2008; 92: 319–341.
7. Bourguignon M. Poisson-geometric INAR(1) process for modelling count time series with overdispersion, Stat. Neerl. 2016; 70(3): 176–192.
8. Weiß CH. Controlling jumps in correlated processes of Poisson counts. Appl Stoch Models Bus Ind. 2009; 25(5): 551–564.
9. Weiß CH. Modelling time series of counts with overdispersion. Stat Methods Appl. 2009; 18(4): 507–519.
10. Jazi MA, Jones G, Lai CD. First-order integer valued AR processes with zero inflated poisson innovations. J Time Ser Anal. 2012; 33(12): 954–963.
11. Jazi MA, Jones G, Lai CD. Integer valued AR(1) with geometric innovations. JIRSS. 2012; 11(2): 173–190.
12. Barreto-Souza W. Zero-modified geometric INAR(1) process for modelling count time series with deflation or inflation of zeros. J Time Ser Anal. 2015; 36(6): 839–852.
13. Bourguignon M. Vasconcellos KLP. First order non-negative integer valued autoregressive processes with power series innovations. Braz J Probab Stat. 2015; 29(1): 71-93.
14. Jose KK, Mariyamma KD. A note on an integer valued time series model with Poisson–negative binomial marginal distribution. Commun Stat Theory Methods. 2016; 45(1): 123-131.
15. Fernández-Fontelo A, Fontdecaba S, Alba A, Puig P. Integer-valued AR processes with hermite innovations and time-varying parameters: An application to bovine fallen stock surveillance at a local scale. Stat Modelling. 2017; 17(3): 172-195.
16. Kim H, Lee S. On first-order integer-valued autoregressive process with Katz family innovations, J Stat Comput Simul. 2017; 87(3): 546–562.
17. Bourguignon M, Weiß CH. An INAR(1) process for modelling count time series with equidispersion, underdispersion and overdispersion. Test. 2017; 26(4): 847–868.
18. Bourguignon M, Rodrigues J, Santos-Neto M. Extended Poisson INAR(1) processes with equidispersion, underdispersion and overdispersion, J Appl Stat. 2019; 46(1): 101-118.
19. Barreto-Souza W. Mixed Poisson INAR(1) processes. Stat Pap. 2019; 60(3): 2119-2139.
20. Delaporte P. Quelques problémes de statistiquemathématique posés par Ìassurance automobile et le bonus non sinistre. Bulletin Trimestriel de l'lnstitut des Actuaires Franąuis. 1959; 227: 87–102.
21. Willmot GE, Sundt B. On evaluation of the Delaporte distribution and related distributions. Scand Actuar J. 1989; 101–113.

22. Karlsen H, Tjøstheim D. Consistent estimates for the NEAR(2) and NLAR(2) time series models. J R Stat Soc Series B Stat Methodol. 1988; 50:313–320.
23. Klimko LA, Nelson PI. On conditional least squares estimation for stochastic processes, Ann Stat. 1978; 6(3): 629–642.
24. Patriota B, Bourguignon M, Santos-Neto M. Tsinteger: Time series of counts analysis, R package version 0.1. 2017.
25. Borges P, Bourguignon M, Molinares FF. A generalized NGINAR(1) process with inflated-parameter geometric counting series. Aust N Z J Stat. 2017; 59(1): 137–150.