مجلهی پژوهشهای آماری ایران

J. Statist. Res. Iran **16** (2019): 287–317

دورهی ۱۶، شمارهی ۲، پاییز و زمستان ۱۳۹۸، صص ۲۸۷–۳۱۷

DOI: 10.52547/jsri.16.2.287

DOR: 20.1001.1.25385771.2020.16.2.3.3

A New Integer-Valued AR(1) Process Based on Power Series Thinning Operator

Eisa Mahmoudi*, Ameneh Rostami and Rasool Roozegar

Yazd University

Received: 2020/17/10 Approved: 2021/22/02

Abstract. In this paper, we introduce the first-order non-negative integer-valued autoregressive (INAR(1)) process with Poisson-Lindley innovations based on a new thinning operator called power series thinning operator. Some statistical properties of process are given. The unknown parameters of the model are estimated by three methods; the conditional least squares, Yule-Walker and conditional maximum likelihood. Then, the performance of these estimators are evaluated using simulation study. Three special cases of model are investigated in some detail. Finally, the model is applied to four real data sets, such as the annual number of earthquakes, the monthly number of measles cases, the numbers of sudden death series and weekly counts of the incidence of acute febrile muco-cutaneous lymph node syndrome. Then we show the potentiality of the model.

Keywords. Integer-valued autoregressive processes; power series distributions; Poisson-Lindley distribution; thinning operator; Yule-Walker equations.

MSC 2010: 62M10.

* Corresponding author

1 Introduction

Sometimes, time series in practice have a discrete nature, such as the number of daily accident on the roads, the number of reserved rooms at a hotel for several days, the number of accidents on a freeway every day, the number of chromosomes interchanges in cells, the number foggy days, the number of bases of DNA sequences, and so on. Count time series are the common type of discrete-valued time series, therefore, the study and analysis of count time series is important and motivates a novel research branch with many practical applications.

One important characteristic of the count time series is the overdispersion property. Until now, for modeling this type of time series, many models have been proposed by various researchers. Ristić et al. (2009) introduced an INAR(1) process based on the negative binomial thinning operator. Aghababaei Jazi et al. (2012) proposed the INAR(1) process with geometric innovations based on the binomial thinning operator. The INAR(1) process with compound-Poisson innovations based on the binomial thinning operator is introduced by Schweer and Weiß (2014). Bourguignon and Vasconcellos (2015) proposed an INAR(1) process with power series innovations based on the binomial thinning operator. Mohammadpour et al. (2018) discussed an INAR(1) process with Poisson-Lindley marginals based on the binomial thinning operator.Lívio et al. (2018) considered the INAR(1) model with Poisson-Lindley innovations based on the binomial thinning operator.

Count data with overdispersion are very frequent in biological, medical sciences, ecology and genetics. For example, the annual number of mothers of the rural area having at least one live birth and one neonatal death, the daily number of hospitalized patients in a hospital, the monthly number of polio cases. Given that, the Poisson-Lindley distribution has many applications in these sciences, therefore, the main idea of this paper is to introduce a INAR(1) process based on this distribution for modeling this type of data.

In this paper, we propose a stationary INAR(1) process with Poisson-Lindley innovations based on the power series thinning operator. In this model, we use the power series thinning operator, which contains, as particular cases, several important and public operators, such as the binomial, negative binomial, Poisson thinning operators. Advantages of the new process are these: (i) this process is appropriate for modeling overdispersed count time series and contains the INARPL(1) model Lívio et al. (2018) as a particular case, (ii) in this process, we use of innovations that come from the Poisson-Lindley distribution that is unimodality, overdispersion, infinite divisibility.

The paper is outlined as follows. In Section 2, we define the power series thinning operator and its properties. In Section 3, we introduce a new stationary first-order integer-valued autoregressive process with Poisson-Lindley innovations based on the power series thinning operator. Some statistical properties of the new process are outlined in this section. Estimation methods for the model parameters are proposed in Section 4. In Section 5, three special cases of the proposed model are studied. Monte Carlo simulation studies is presented to compare the behaviors of the estimators in Section 6. In Section 7, we provide applications to four real data sets and discuss the obtained results. Finally, Section 8 concludes the paper.

2 The Power Series Thinning Operator

In this section, we introduce the power series (PS) thinning operator. Therefore, we first review the PS family of distributions and focus on some properties of this family. The random variable Y with probability mass function

$$P(Y=y)=\frac{a(y)\beta^y}{C(\beta)}, \qquad y\in T, \tag{1}$$

has PS distribution with range T, where T is a subset of the non-negative integer numbers and $C(\beta)=\sum_T a(y)\beta^y$ is finite for all $\beta\in(0,t)$, $a(y)>0$. This family of distributions includes some well-known distributions, such as Bernoulli, binomial, Poisson, geometric, negative binomial and logarithmic distribution.

The expectation, variance and probability generating function (pgf) of Y are given, respectively, by

$$\begin{aligned} E(Y) &= \beta G^{'}(\beta), \\ Var(Y) &= \beta G^{'}(\beta)+\beta^2 G^{''}(\beta), \\ \varphi_Y(s) &= \frac{C(s\beta)}{C(\beta)}, \end{aligned}$$

where $G(\beta)=\log C(\beta)$, $G^{'}(\beta)=\frac{d}{d\beta}G(\beta)$ and $G^{''}(\beta)=\frac{d^2}{d\beta^2}G(\beta)$.

In the next Proposition, we prove the PS family of distribution are closed relative to sums of independent variables.

Proposition 1. Suppose that $Y_1, Y_2, ..., Y_n$ is a sequence of independent and identically distributed (i.i.d) PS random variables, relative to the function C and with parameter value $\beta \in (0, t)$. Then $Y_1 + Y_2 + ... + Y_n$ has the PS distribution relative to the function C^n and with parameter β.

Proof. For proof, we use of pgf of the sums of independent variables.

$$\begin{aligned} \varphi_{Y_1+Y_2+...+Y_n}(s) &= \varphi_{Y_1}(s)\varphi_{Y_2}(s)...\varphi_{Y_n}(s) \\ &= \frac{C(s\beta)}{C(\beta)}\frac{C(s\beta)}{C(\beta)} \cdot \frac{C(s\beta)}{C(\beta)} = (\frac{C(s\beta)}{C(\beta)})^n, \qquad s < \frac{t}{\beta}. \end{aligned}$$

□

Definition 1. (PS thinning operator)

Let X be a non-negative integer-valued random variable and $\alpha > 0$. The PS thinning operator, denoted by αo_{ps}, is defined by

$$\alpha o_{ps} X = \sum_{i=1}^{X} Y_i, \tag{2}$$

where counting series $\{Y_i\}$ is a sequence of independent and identically distributed PS random variables with the mean α and the variance δ, mutually independent of X. Note that for $X = 0$, the empty sum is defined as 0.

Table 1 presents the parameter β, δ and functions $C(s\beta)$ and $a(y)$ corresponding to some special cases of the PS family of distribution with mean α. Note that, the logarithmic series distribution is not considered in the paper, because the parameter β in this distribution does not have an explicit expression.

Remark 1. Special cases of the PS thinning operator are:

(i) If $\beta = \frac{\alpha}{1-\alpha}$ and $C(\beta) = \frac{1}{1-\alpha}$ then the PS thinning operator becomes the binomial thinning operator proposed by Steutel and van Harn (1979).

(ii) If $\beta = \frac{\alpha}{1+\alpha}$ and $C(\beta) = 1+\alpha$ then the PS thinning operator becomes the negative binomial thinning operator proposed by Ristić et al. (2009).

Table 1. Some special cases of the PS family.

Distribution	β	δ	$C(\beta)$	$C(s\beta)$	$a(y)$	T	t
Bernoulli	$\frac{\alpha}{1-\alpha}$	$\alpha(1-\alpha)$	$\frac{1}{1-\alpha}$	$\frac{1-\alpha+\alpha s}{1-\alpha}$	1	$\{0,1\}$	∞
Binomial	$\frac{\frac{\alpha}{n}}{1-\frac{\alpha}{n}}$	$n\alpha(1-\frac{\alpha}{n})$	$(\frac{1}{1-\frac{\alpha}{n}})^n$	$(\frac{1-\frac{\alpha}{n}+\frac{\alpha s}{n}}{1-\frac{\alpha}{n}})^n$	$\binom{n}{y}$	$\{0,1,2,...,n\}$	∞
Poisson	α	α	e^{α}	$e^{\alpha s}$	$\frac{1}{y!}$	$\{0,1,...\}$	∞
Geometric	$\frac{\alpha}{1+\alpha}$	$\alpha(1+\alpha)$	$1+\alpha$	$\frac{1+\alpha}{1+\alpha-\alpha s}$	1	$\{0,1,...\}$	1
Negative Binomial	$\frac{\alpha}{r+\alpha}$	$\frac{\alpha(r+\alpha)}{r}$	$(\frac{r+\alpha}{r})^r$	$(\frac{r+\alpha}{r+\alpha-\alpha s})^r$	$\frac{\Gamma(r+y)}{y!\Gamma(r)}$	$\{0,1,...\}$	1

(iii) If $\beta=\alpha$ and $C(\beta)=e^{\alpha}$ then the PS thinning operator becomes the Poisson thinning operator proposed by Ferland et al. (2006).

From Proposition 1 and Definition 1, we have that, for a given random variable X, $\alpha o_{ps} X$ in (2) has a PS distribution relative to the function C^X and with parameter β. This implies that the pgf of $\alpha o_{ps} X|X$ is given as

$$\varphi_{\alpha o_{ps} X|X}(s)=(\frac{C(s\beta)}{C(\beta)})^X, \qquad s<\frac{t}{\beta}.$$

Remark 2. (i) If $\beta=\frac{\alpha}{1-\alpha}$ and $C(\beta)=\frac{1}{1-\alpha}$, then $\alpha o_{ps} X|X$ has a binomial distribution with parameters X and α, its probability mass function (pmf) is $P(\alpha o_{ps} X=m|X)=\binom{X}{m}\alpha^m(1-\alpha)^{X-m}$.

(ii) If $\beta=\frac{\alpha}{1+\alpha}$ and $C(\beta)=1+\alpha$, then $\alpha o_{ps} X|X$ has a negative binomial distribution with parameters X and $\frac{1}{1+\alpha}$, its pmf is $P(\alpha o_{ps} X=m|X)=\binom{X+m-1}{m}\left(\frac{1}{1+\alpha}\right)^X\left(\frac{\alpha}{1+\alpha}\right)^m$.

(iii) If $\beta=\alpha$ and $C(\beta)=e^{\alpha}$, then $\alpha o_{ps} X|X$ has a Poisson distribution with parameters αX, its pmf is $P(\alpha o_{ps} X=m|X)=\frac{e^{-\alpha X}(\alpha X)^m}{m!}$.

(iv) If $\beta=\frac{\frac{\alpha}{n}}{1-\frac{\alpha}{n}}$ and $C(\beta)=(\frac{1}{1-\frac{\alpha}{n}})^n$, then $\alpha o_{ps} X|X$ has a binomial distribution with parameters nX and α, its pmf is $P(\alpha o_{ps} X=m|X)=\binom{nX}{m}\alpha^m(1-\alpha)^{nX-m}$.

(v) If $\beta=\frac{\alpha}{r+\alpha}$ and $C(\beta)=(\frac{r+\alpha}{r})^r$, then $\alpha o_{ps} X|X$ has a negative binomial distribution with parameters rX and $\frac{1}{1+\alpha}$, its pmf is $P(\alpha o_{ps} X=m|X)=\binom{rX+m-1}{m}\left(\frac{1}{1+\alpha}\right)^{rX}\left(\frac{\alpha}{1+\alpha}\right)^m$.

Proposition 1. The PS thinning operator has the following properties:

(i) $E(\alpha o_{ps} X|X) = \alpha X$

(ii) $Var(\alpha o_{ps} X|X) = \delta X$

(iii) $E(\alpha o_{ps} X) = \alpha E(X)$

(iv) $Var(\alpha o_{ps} X) = \alpha^2 Var(X) + \delta E(X)$

Proof. (i)

$$\begin{aligned}
E(\alpha o_{ps} X|X) =& \varphi'_{\alpha o_{ps} X|X}(s)|_{s=1} \\
=& \frac{d}{ds}(\frac{C(s\beta)}{C(\beta)})^X|_{s=1} = \frac{X\beta C'(s\beta)}{C(\beta)}(\frac{C(s\beta)}{C(\beta)})^{X-1}|_{s=1} \\
=& \frac{X\beta C'(\beta)}{C(\beta)}(\frac{C(\beta)}{C(\beta)})^{X-1} = \frac{X\beta C'(\beta)}{C(\beta)} \\
=& X\beta G'(\beta) = XE(Y) = X\alpha.
\end{aligned}$$

(ii)

$$\begin{aligned}
Var(\alpha o_{ps} X|X) =& E((\alpha o_{ps} X)^2|X) - E^2(\alpha o_{ps} X|X) \\
=& \varphi''_{\alpha o_{ps} X|X}(s)|_{s=1} + E(\alpha o_{ps} X|X) - E^2(\alpha o_{ps} X|X) \\
=& \frac{d^2}{ds^2}(\frac{C(s\beta)}{C(\beta)})^X|_{s=1} + XE(Y) - X^2E^2(Y) \\
=& (\frac{X(X-1)\beta^2 C'^2(s\beta)}{C^2(\beta)}(\frac{C(s\beta)}{C(\beta)})^{X-2} \\
&+ \frac{X\beta^2 C''(s\beta)}{C(\beta)}(\frac{C(s\beta)}{C(\beta)})^{X-1})|_{s=1} + XE(Y) - X^2E^2(Y) \\
=& \frac{X(X-1)\beta^2 C'^2(s\beta) + X\beta^2 C(\beta)C''(s\beta)}{C^2(\beta)} \\
&+ XE(Y) - X^2E^2(Y) \\
=& \frac{X^2\beta^2 C'^2(\beta)}{C^2(\beta)} + X(Var(Y) - \frac{\beta C'(\beta)C(\beta)}{C^2(\beta)}) \\
&+ XE(Y) - X^2E^2(Y) \\
=& X^2E^2(Y) + XVar(Y) - X^2E^2(Y) = XVar(Y) = X\delta.
\end{aligned}$$

(iii)

$$E(\alpha o_{ps} X) = E(E(\alpha o_{ps} X|X)) = E(\alpha X) = \alpha E(X).$$

(iv)

$$\begin{aligned} Var(\alpha o_{ps} X) =& Var(E(\alpha o_{ps} X|X)) + E(Var(\alpha o_{ps} X|X)) \\ =& Var(\alpha X) + E(\delta X) = \alpha^2 Var(X) + \delta E(X). \end{aligned}$$

□

3 Construction of the Model

In this section, we introduce a stationary INAR(1) process generated by the PS thinning operator with Poisson-Lindley (PL) innovations (PSINARPL(1)). First, we give a brief peruse of the PL distribution.

A random variable W is said to have a PL distribution with parameter θ if its pmf is of the form

$$P(W = w) = \frac{\theta^2(w + \theta + 2)}{(\theta + 1)^{w+3}}, \quad w = 0, 1, 2, ... \qquad \theta > 0,$$

which was introduced firstly by Sankaran Sankaran (1970). The mean and variance of W are given by

$$E(W) = \frac{\theta + 2}{\theta(\theta + 1)}, \qquad Var(W) = \frac{\theta^3 + 4\theta^2 + 6\theta + 2}{\theta^2(\theta + 1)^2}.$$

The dispersion index is given by

$$I_w = \frac{\theta^3 + 4\theta^2 + 6\theta + 2}{\theta^3 + 3\theta^2 + 2\theta} > 1,$$

that is, the PL distribution is overdispersion. The pgf and moment generat-

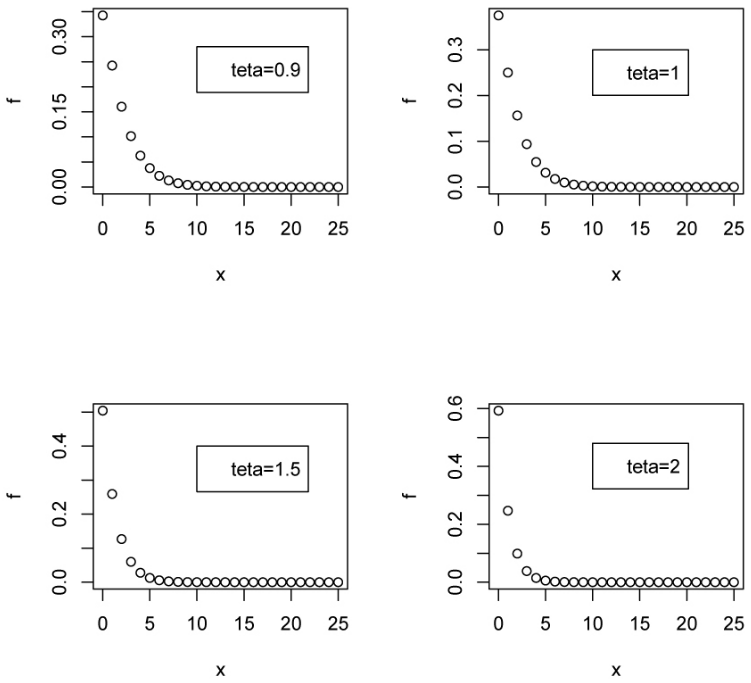

Figure 1. Probability mass function of PL distribution for different values θ.

ing function (mgf) are

$$\varphi_W(s) = \frac{\theta^2}{1+\theta}\left[\frac{1}{(1+\theta-s)^2} + \frac{1}{(1+\theta-s)}\right],$$
$$M_W(s) = \frac{\theta^2}{1+\theta}\left[\frac{1}{(1+\theta-e^s)^2} + \frac{1}{(1+\theta-e^s)}\right].$$

Figure 1 shows the pmf of the PL distribution for different values θ.

Definition 2. (PSINARPL(1) process) The PSINARPL(1) model is defined as

$$X_t = \alpha o_{ps} X_{t-1} + W_t, \qquad t \geq 1, \tag{3}$$

where $\alpha \in [0,1)$ and W_t's are independent and identically distributed random variables from PL distribution that are independent from Y_i's and X_{t-l} for $l \geq 1$.

3.1 Statistical Properties of the Model

The mean, variance and index of dispersion of X_t are given, respectively, by

$$\begin{aligned} E(X_t) &= \frac{\theta+2}{\theta(\theta+1)(1-\alpha)}, \\ Var(X_t) &= \frac{\delta(\theta+2)}{\theta(\theta+1)(1-\alpha)(1-\alpha^2)} + \frac{\theta^3+4\theta^2+6\theta+2}{\theta^2(\theta+1)^2(1-\alpha^2)}, \\ I_{X_t} &= \frac{\theta(\theta+1)(\theta+2)\delta+(1-\alpha)(\theta^3+4\theta^2+6\theta+2)}{\theta(\theta+1)(\theta+2)(1-\alpha^2)}. \end{aligned}$$

Since $I_{X_t} > 1$, thus, the PSINARPL(1) process is overdispersion. The conditional pgf of X_t given X_{t-1} becomes

$$\varphi_{X_t|X_{t-1}}(s) = (\frac{C(s\beta)}{C(\beta)})^{X_{t-1}}(\frac{\theta^2}{1+\theta}[\frac{1}{(1+\theta-s)^2} + \frac{1}{(1+\theta-s)}]).$$

The autocovariance and autocorrelation functions for model (2) are, respectively

$$\gamma_k = Cov(X_{t-k}, X_t) = \alpha^k\gamma_0, \qquad \rho_k = \frac{\gamma_k}{\gamma_0} = \frac{a^k\gamma_0}{\gamma_0} = \alpha^k.$$

According to Al-Osh and Alzaid (1987) and Theorem 2.1 from Du and Li (1991), the process X_t satisfying (3) is second-order stationary if $0 \leq \alpha < 1$. In the case of $\alpha = 1$ the process is non-stationary, because in this case the mean and variance of X_t are not defined.

The conditional mean and conditional variance of the proposed model are obtained as follows

$$E(X_{t+1}|X_t) = \alpha X_t + \mu_w, \qquad Var(X_{t+1}|X_t) = \delta X_t + \sigma_w^2,$$

where $E(W_t) = \mu_w$ and $Var(W_t) = \sigma_w^2$.

***Theorem* 1.** *The k-step ahead conditional mean and variance the PSINARPL(1) model are given by*

$$E(X_{t+k}|X_t) = \alpha^k X_t + \mu_w \sum_{i=1}^{k} \alpha^{i-1}, \tag{4}$$

$$Var(X_{t+k}|X_t) = \alpha^{2k}X_t^2 + (\sigma_w^2 + \mu_w^2)\sum_{i=1}^{k}\alpha^{2i-2} + (2\alpha\mu_w + \delta)(X_t\sum_{i=k}^{2k-1}\alpha^{i-1} + \mu_w(\sum_{j=0}^{k-1}\sum_{i=j}^{2j}\alpha^i)) - (\alpha^k X_t + \mu_w\sum_{i=1}^{k}\alpha^{i-1})^2, \qquad k \geq 2. \tag{5}$$

respectively.

Proof. For the first part since

$$\begin{aligned} E(X_{t+1}|X_t) = E(\alpha o_{ps} X_t + W_{t+1}|X_t) &= E(\alpha o_{ps} X_t|X_t) + E(W_{t+1}) \\ &= \alpha X_t + \mu_w, \end{aligned}$$

and

$$\begin{aligned} E(X_{t+2}|X_t) = E(\alpha o_{ps} X_{t+1} + W_{t+2}|X_t) &= E(E(\alpha o_{ps} X_{t+1}|X_{t+1})|X_t) + \mu_w \\ &= E(\alpha X_{t+1}|X_t) + \mu_w \\ &= \alpha E(X_{t+1}|X_t) + \mu_w \\ &= \alpha^2 X_t + (1+\alpha)\mu_w, \end{aligned}$$

hence by induction, we have

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

For the proof of the second part, we use the definition of the conditional variance

$$Var(X_{t+k}|X_t) = E(X_{t+k}^2|X_t) - E^2(X_{t+k}|X_t), \tag{6}$$

now we obtain an expression for $E(X^2_{t+k}|X_t)$, since

$$\begin{aligned} E(X^2_{t+1}|X_t) &= E((\alpha o_{ps} X_t)^2|X_t) + E(W^2_{t+1}) + 2E(W_{t+1})E(\alpha o_{ps} X_t|X_t) \\ &= \delta X_t + \alpha^2 X^2_t + \sigma^2_w + \mu^2_w + 2\alpha\mu_w X_t, \end{aligned}$$

and

$$\begin{aligned} E(X^2_{t+2}|X_t) &= E((\alpha o_{ps} X_{t+1})^2|X_t) + E(W^2_{t+2}) + 2E(W_{t+2})E(\alpha o_{ps} X_{t+1}|X_t) \\ &= E(E(\alpha o_{ps} X_{t+1})^2|X_{t+1})|X_t) + \sigma^2_w + \mu^2_w \\ &+ 2\mu_w E(E(\alpha o_{ps} X_{t+1}|X_{t+1})|X_t) \\ &= E(\delta X_{t+1} + \alpha^2 X^2_{t+1}|X_t) + \sigma^2_w + \mu^2_w + 2\mu_w E(\alpha X_{t+1}|X_t) \\ &= \alpha^4 X^2_t + (\sigma^2_w + \mu^2_w)(1+\alpha^2) + (2\alpha\mu_w + \delta)((\alpha+\alpha^2)X_t + \mu_w). \end{aligned}$$

hence

$$\begin{aligned} E(X^2_{t+k}|X_t) = \alpha^{2k} X^2_t + (\sigma^2_w + \mu^2_w)\sum_{i=1}^{k} \alpha^{2i-2} + (2\alpha\mu_w + \delta)(X_t \sum_{i=k}^{2k-1} \alpha^{i-1} \\ + \mu_w(\sum_{j=0}^{k-1}\sum_{i=j}^{2j} \alpha^i)), \qquad k \geq 2. \end{aligned} \tag{7}$$

Thus, the k-step ahead conditional variance is achieved with replacing (4) and (7) in Equation (6). □

Note that

$$\begin{aligned} \lim_{k\to\infty} E(X_{t+k}|X_t) &= \frac{\mu_w}{1-\alpha}, \\ \lim_{k\to\infty} Var(X_{t+k}|X_t) &= \frac{\sigma^2_w + \mu^2_w}{1-\alpha^2} + \frac{(2\mu_w\alpha + \delta)\mu_w}{(1-\alpha^2)(1-\alpha)} - \frac{\mu^2_w}{(1-\alpha)^2} \\ &= \frac{\delta\mu_w}{(1-\alpha)(1-\alpha^2)} + \frac{\sigma^2_w}{1-\alpha^2} \end{aligned}$$

which are the unconditional mean and variance of the process.

The transition probabilities of the model are given by

$$\begin{aligned} P_{lk} &= P(X_t = k|X_{t-1} = l) = P(\alpha o_{ps} X_{t-1} + W_t = k|X_{t-1} = l) \\ &= \sum_m P(\alpha o_{ps} X_{t-1} = m|X_{t-1} = l) P(W_t = k - m). \end{aligned}$$

If $T = \{0, 1, \cdots, l\}$ for fixed $l \in Z^+$ then $0 \le m \le l$ and $k - m \ge 0$ implies $0 \le m \le \min(l, k)$, so we have

$$P_{lk} = \sum_{m=0}^{\min(l,k)} P(\alpha o_{ps} X_{t-1} = m|X_{t-1} = l) \left[\frac{\theta^2((k-m) + \theta + 2)}{(\theta+1)^{k-m+3}} I_{\{0,1,\ldots\}}(k-m) \right].$$

If $T = \{0, 1, 2, \cdots\}$ then $m \ge 0$ and $k - m \ge 0 \Longrightarrow 0 \le m \le k$, so

$$\begin{aligned} P_{lk} &= \sum_{m=0}^{k} P(\alpha o_{ps} X_{t-1} = m|X_{t-1} = l) \left[\frac{\theta^2(k-m+\theta+2)}{(\theta+1)^{k-m+3}} I_{\{0,1,\ldots\}}(k-m) \right] \\ &\times I(l \neq 0) + \left[\frac{\theta^2(k+\theta+2)}{(\theta+1)^{k+3}} I_{\{0,1,\ldots\}}(k) \right] I(l = 0). \end{aligned}$$

Where $P(\alpha o_{ps} X_{t-1} = m|X_{t-1} = l)$ is given in Remark 2.

4 Estimation of the Unknown Parameters

In this section, we derive the estimators of the unknown parameters using the Yule-Walker, conditional least squares and conditional maximum likelihood methods. Note that, for obtain the estimator of θ, we use the auxiliary parameter μ. Let $X_1, ..., X_T, T \in N$ be a random sample of size T from the PSINARPL(1) process.

4.1 The Conditional Least Squares (CLS) Method

The CLS estimators of the parameters α and μ for the PSINARPL(1) model is obtained by minimizing the following function:

$$S_n(\alpha, \mu) = \sum_{t=2}^{T} (X_t - E(X_t|X_{t-1}))^2 = \sum_{t=2}^{T} (X_t - \alpha X_{t-1} - (1-\alpha)\mu)^2,$$

where $\mu = E(X_t)$. Thus, the conditional least squares estimator of α and μ are given as follow,

$$\hat{\alpha}_{CLS} = \frac{(T-1)\sum_{t=2}^{T} X_t X_{t-1} - \sum_{t=2}^{T} X_t \sum_{t=2}^{T} X_{t-1}}{(T-1)\sum_{t=2}^{T} X_{t-1}^2 - \left(\sum_{t=2}^{T} X_{t-1}\right)^2},$$

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

Also, the estimator for θ is obtained by solving the equation

$$\hat{\mu}_{CLS} = \frac{\theta+2}{\theta(\theta+1)(1-\hat{\alpha}_{CLS})}.$$

So, $\hat{\theta}_{CLS}$ is given by

$$\hat{\theta}_{CLS} = \frac{(1-(1-\hat{\alpha}_{CLS})\hat{\mu}_{CLS}) + \sqrt{((1-\hat{\alpha}_{CLS})\hat{\mu}_{CLS}-1)^2 + 8(1-\hat{\alpha}_{CLS})\hat{\mu}_{CLS}}}{2(1-\hat{\alpha}_{CLS})\hat{\mu}_{CLS}}.$$

Theorem 2. *The estimators $\hat{\alpha}_{CLS}$ and $\hat{\theta}_{CLS}$ are strongly consistent for estimating α and θ, respectively, and satisfy the asymptotic normality*

$$\sqrt{n}\begin{pmatrix}\hat{\alpha}_{CLS}-\alpha \\ \hat{\theta}_{CLS}-\theta\end{pmatrix} \to^d N\left(0, c^2 A\right), \tag{8}$$

where

$$A = \begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{pmatrix},$$

and

$$c = \theta^2(\theta+1)^2[(\mu_2-\mu_1^2)(\theta^2+4\theta+2)]^{-1},$$

$$r_{11} = \frac{(\theta^2+4\theta+2)^2}{\theta^4(\theta+1)^4}[(\delta\mu_3+\mu_2\sigma_W^2) - \mu_1(\delta\mu_2+\mu_1\sigma_W^2)$$
$$+\mu_1(\mu_1(\delta\mu_1+\sigma_W^2) - (\delta\mu_2+\mu_1\sigma_W^2))],$$

$$r_{12} = r_{21} = \frac{(\theta^2+4\theta+2)}{\theta^2(\theta+1)^2}[\mu_1(\delta\mu_3+\sigma_W^2\mu_2)-\mu_2(\delta\mu_2+\mu_1\sigma_W^2) + \mu_1\mu_2(\delta\mu_1+\sigma_W^2)-\mu_1^2(\delta\mu_2+\mu_1\sigma_W^2)],$$

$$r_{22} = \mu_1^2(\delta\mu_3+\sigma_W^2\mu_2)-2\mu_1\mu_2(\delta\mu_2+\mu_1\sigma_W^2)+\mu_2^2(\delta\mu_1+\sigma_W^2),$$

$E(X_t^r)=\mu_r, \quad r=1,2,3$

Proof. The proof is similar to proof of Theorem 3.1 of Lívio et al. (2018), using this fact that in the proof we will consider $f_{t|t-1}=\delta X_t+\sigma_W^2$. So the proof is omitted. □

4.2 The Yule-Walker (YW) Method

Since $\mu = E(X_t)$ and $\alpha = \frac{\gamma(1)}{\gamma(0)}$, then in model refdefn2.2, the YW estimators of α and μ are obtained as follows

$$\hat{\mu}_{YW} = \bar{X}_T = \frac{1}{T}\sum_{t=1}^{T} X_t,$$

$$\hat{\alpha}_{YW} = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)} = \frac{\sum_{t=2}^{T}(X_t-\bar{X}_t)((X_{t-1}-\bar{X}_t)}{\sum_{t=1}^{T}(X_t-\bar{X}_t)^2}.$$

The YW estimator of θ is given by

$$\hat{\theta}_{YW} = \frac{(1-(1-\hat{\alpha}_{YW})\hat{\mu}_{YW})+\sqrt{((1-\hat{\alpha}_{YW})\hat{\mu}_{YW}-1)^2+8(1-\hat{\alpha}_{YW})\hat{\mu}_{YW}}}{2(1-\hat{\alpha}_{YW})\hat{\mu}_{YW}}.$$

4.3 The Conditional Maximum Likelihood (CML) Method

The ML estimators of α and θ are obtained by maximizing the likelihood function

$$L(\theta,\alpha|\mathbf{x}) = f(x_1,...,x_n) = f(x_1)f(x_2|x_1)...f(x_n|x_{n-1}).$$

In general case, obtaining the marginal distribution of X_1 is hard, therefore, a simple method is to condition on X_1. Thus, we use the CML method.

The CML estimators of α and θ in the model are obtained by maximizes the conditional log-likelihood function. The CML estimators no closed form, so, in practice, these estimators are obtained using numerical methods.

5 Special Sases of the PSINARPL(1) Process

In this section, we consider three special cases of the PSINARPL(1) process and studied some properties of the models.

5.1 INARPL(1) Process Based on the Binomial Thinning Operator

If $\beta = \frac{\alpha}{1-\alpha}$ and $C(\beta) = \frac{1}{1-\alpha}$, X_t is the stationary INARPL(1) process based on the binomial thinning operator introduced by Lívio et al. (2018).
The mean and the variance of X_t are given by

$$\begin{aligned} E(X_t) &= \frac{\theta+2}{\theta(\theta+1)(1-\alpha)}, \\ Var(X_t) &= \frac{\alpha(\theta+2)}{\theta(\theta+1)(1-\alpha^2)} + \frac{\theta^3+4\theta^2+6\theta+2}{\theta^2(\theta+1)^2(1-\alpha^2)}. \end{aligned}$$

The conditional expectation and variance of X_t are

$$\begin{aligned} E(X_{t+1}|X_t) &= \alpha X_t + \frac{\theta+2}{\theta(\theta+1)}, \\ Var(X_{t+1}|X_t) &= \alpha(1-\alpha)X_t + \frac{\theta^3+4\theta^2+6\theta+2}{\theta^2(\theta+1)^2}. \end{aligned}$$

The transition probabilities of the BINARPL(1) model are given by

$$P_{lk} = \sum_{m=0}^{\min(l,k)} \binom{l}{m} \alpha^m (1-\alpha)^{l-m} \left[\frac{\theta^2((k-m)+\theta+2)}{(\theta+1)^{k-m+3}} I_{\{0,1,\dots\}}(k-m) \right].$$

5.2 INARPL(1) Based on the Negative Binomial Thinning Operator

If $\beta = \frac{\alpha}{1+\alpha}$ and $C(\beta) = 1+\alpha$, X_t is the stationary INARPL(1) process based on the negative binomial thinning operator.

The mean and the variance of X_t are given by

$$\begin{aligned} E(X_t) &= \frac{\theta+2}{\theta(\theta+1)(1-\alpha)}, \\ Var(X_t) &= \frac{\alpha(\theta+2)}{\theta(\theta+1)(1-\alpha)^2} + \frac{\theta^3+4\theta^2+6\theta+2}{\theta^2(\theta+1)^2(1-\alpha^2)}. \end{aligned}$$

The conditional expectation and variance of X_t are

$$\begin{aligned} E(X_{t+1}|X_t) &= \alpha X_t + \frac{\theta+2}{\theta(\theta+1)}, \\ Var(X_{t+1}|X_t) &= \alpha(1+\alpha)X_t + \frac{\theta^3+4\theta^2+6\theta+2}{\theta^2(\theta+1)^2}. \end{aligned}$$

The transition probabilities of the NBINARPL(1) model are given by

$$\begin{aligned} P_{lk} = &\sum_{m=0}^{k} \binom{l+m-1}{m} \left(\frac{1}{1+\alpha}\right)^l \left(\frac{\alpha}{1+\alpha}\right)^m \left[\frac{\theta^2(k-m+\theta+2)}{(\theta+1)^{k-m+3}} I_{\{0,1,\ldots\}}(k-m)\right] \\ &\times I(l \neq 0) + \left[\frac{\theta^2(k+\theta+2)}{(\theta+1)^{k+3}} I_{\{0,1,\ldots\}}(k)\right] I(l=0). \end{aligned}$$

5.3 INARPL(1) Based on the Poisson Thinning Operator

If $\beta = \alpha$ and $C(\beta) = e^{\alpha}$, X_t is the stationary INARPL(1) process based on the Poisson thinning operator.

The mean and variance of X_t are

$$\begin{aligned} E(X_t) &= \frac{\theta+2}{\theta(\theta+1)(1-\alpha)}, \\ Var(X_t) &= \frac{\alpha(\theta+2)}{\theta(\theta+1)(1-\alpha)(1-\alpha^2)} + \frac{\theta^3+4\theta^2+6\theta+2}{\theta^2(\theta+1)^2(1-\alpha^2)}. \end{aligned}$$

The conditional expectation and variance $\beta = \alpha$ are given by

$$\begin{aligned} E(X_{t+1}|X_t) &= \alpha X_t + \frac{\theta+2}{\theta(\theta+1)}, \\ Var(X_{t+1}|X_t) &= \alpha X_t + \frac{\theta^3+4\theta^2+6\theta+2}{\theta^2(\theta+1)^2}. \end{aligned}$$

Also, the transition probabilities of the PINARPL(1) model are given by

$$P_{lk} = \sum_{m=0}^{k} \frac{e^{-\alpha l}(\alpha l)^m}{m!} \left[\frac{\theta^2(k-m+\theta+2)}{(\theta+1)^{k-m+3}} I_{\{0,1,\ldots\}}(k-m) \right] \times I(l \neq 0) + \left[\frac{\theta^2(k+\theta+2)}{(\theta+1)^{k+3}} I_{\{0,1,\ldots\}}(k) \right] I(l=0).$$

6 Some Numerical Results

In this section, for each three models, BINARPL(1), NBINARPL(1) and PINARPL(1), we produce 1000 samples of size T = 100, 200 and 300 and obtain the estimates of the parameters using three methods presented in section 4, then, we compare the performance of the estimators.
Tables 2, 3 and 4 present the empirical mean (EM), bias and root mean square error (RMSE) of the estimates of the parameters in each three proposed models. As we can see from those tables, the estimators converge to the true value and the RMSE decrease as the size of the sample increases.
In two models, BINARPL(1) and NBINARPL(1), the RMSE of the CML estimators are less than the others. In the PINARPL(1) model, RMSE of the CLS and YW estimators of α are less than the CML estimator while the CML estimator of θ have the smallest RMSE than the others.

7 Real Data Examples

In this section, we fit PSINARPL(1) model to four real data sets and show the usefulness of the model. For this purpose, we compare the BINARPL(1), NBINARPL(1), PINARPL(1) models with the INARP(1) (INAR(1) with Poisson innovations based on the binomial thinning operator) proposed by Al-Osh and Alzaid (1987) and INARG(1) (INAR(1) with geometric innovations based on the binomial thinning operator) proposed by Aghababaei Jazi et al. (2012) models.

7.1 The Number of Earthquakes Magnitude 7.0 or Greater

The first example assumes the number of earthquakes magnitude 7.0 or greater, annually from 1900-1998 in the world. The sample path, autocorrelation and partial autocorrelation functions of series are shown in Figure

Table 2. Empirical means, bias and root mean squared errors of the estimates of α and θ for BINARPL(1).

Sample size	$\hat{\alpha}_{BCLS}$	$\hat{\theta}_{BCLS}$	$\hat{\alpha}_{BYW}$	$\hat{\theta}_{BYW}$	$\hat{\alpha}_{BCML}$	$\hat{\theta}_{BCML}$
True value $\alpha = 0.2$ and $\theta = 0.6$						
T=100						
EM	0.1704	0.5826	0.1687	0.5877	0.2040	0.6099
Bias	-0.0295	-0.0174	-0.0313	-0.0122	0.0040	0.0099
RMSE	0.1017	0.0802	0.1013	0.0799	0.0582	0.0692
T=200						
EM	0.1853	0.5888	0.1845	0.5915	0.1994	0.6026
Bias	-0.0146	-0.0111	-0.0155	-0.0085	-0.0006	0.0026
RMSE	0.0739	0.0591	0.0738	0.0590	0.0415	0.0474
T=300						
EM	0.1831	0.5867	0.1825	0.5884	0.2011	0.6047
Bias	-0.0169	-0.0133	-0.0175	-0.0116	0.0011	0.0047
RMSE	0.0596	0.0473	0.0596	0.0470	0.0361	0.0403
True value $\alpha = 0.5$ and $\theta = 1$						
T=100						
EM	0.4483	0.9227	0.4435	0.9308	0.4980	1.0077
Bias	-0.0517	-0.0772	-0.0565	-0.0692	-0.0020	0.0077
RMSE	0.1089	0.1732	0.1106	0.1713	0.0542	0.1349
T=200						
EM	0.4591	0.9270	0.4570	0.9312	0.4999	1.0138
Bias	-0.0409	-0.0730	-0.0430	-0.0687	-0.0001	0.0138
RMSE	0.0779	0.1315	0.0789	0.1300	0.0395	0.0949
T=300						
EM	0.4629	0.9259	0.4612	0.9284	0.5010	1.0077
Bias	-0.0371	-0.0741	-0.0388	-0.0716	0.0010	0.0077
RMSE	0.0674	0.1178	0.0683	0.1165	0.0313	0.0799
True value $\alpha = 0.9$ and $\theta = 2$						
T=100						
EM	0.8603	1.7063	0.8504	1.7235	0.8974	2.0404
Bias	-0.0397	-0.2937	-0.0495	-0.2764	-0.0026	0.0404
RMSE	0.0664	0.6176	0.0728	0.6245	0.0174	0.3322
T=200						
EM	0.8793	1.8186	0.8743	1.8230	0.8992	2.0273
Bias	-0.0207	-0.1814	-0.0257	-0.1770	-0.0008	0.0273
RMSE	0.0421	0.5074	0.0450	0.5118	0.0113	0.2319
T=300						
EM	0.8853	1.8613	0.8821	1.8657	0.8995	2.0232
Bias	-0.0147	-0.1387	-0.0178	-0.1342	-0.0005	0.0232
RMSE	0.0314	0.4179	0.0330	0.4143	0.0095	0.1902

Table 3. Empirical means, bias and root mean squared errors of the estimates of α and θ for NBINARPL(1).

Sample size	$\hat{\alpha}_{NBCLS}$	$\hat{\theta}_{NBCLS}$	$\hat{\alpha}_{NBYW}$	$\hat{\theta}_{NBYW}$	$\hat{\alpha}_{NBCML}$	$\hat{\theta}_{NBCML}$
True value $\alpha = 0.2$ and $\theta = 0.6$						
T=100						
EM	0.1718	0.5839	0.1702	0.5892	0.1993	0.6084
Bias	-0.0282	-0.0160	-0.0298	-0.0108	-0.0007	0.0084
RMSE	0.1072	0.0831	0.1065	0.0829	0.0761	0.0778
T=200						
EM	0.1840	0.5875	0.1832	0.5902	0.2019	0.6056
Bias	-0.0160	-0.0124	-0.0168	-0.0098	0.0019	0.0056
RMSE	0.07567	0.0610	0.0757	0.0609	0.0535	0.0546
T=300						
EM	0.1832	0.5847	0.1826	0.5864	0.2017	0.6051
Bias	-0.0168	-0.0153	-0.0174	-0.0135	0.0017	0.0051
RMSE	0.0620	0.0482	0.0620	0.0478	0.0446	0.0427
True value α=0.5 and $\theta = 1$						
T=100						
EM	0.4359	0.8859	0.4314	0.8936	0.4793	1.0103
Bias	-0.0641	-0.1141	-0.0686	-0.1064	-0.0207	0.0103
RMSE	0.1216	0.1823	0.1234	0.1785	0.0944	0.1765
T=200						
EM	0.4504	0.9022	0.4480	0.9059	0.4953	1.0077
Bias	-0.0496	-0.0978	-0.0520	-0.0941	-0.0047	0.0077
RMSE	0.0899	0.1485	0.0910	0.1464	0.0641	0.1203
T=300						
EM	0.4584	0.9072	0.4567	0.9096	0.4922	1.0026
Bias	-0.0416	-0.0927	-0.0433	-0.0904	-0.0078	0.0026
RMSE	0.0749	0.1312	0.0757	0.1297	0.0541	0.1021
True value $\alpha = 0.9$ and $\theta = 2$						
T=100						
EM	0.8047	1.3007	0.7935	1.2959	0.8520	1.9061
Bias	-0.0953	-0.6993	-0.1065	-0.7041	-0.0480	-0.0939
RMSE	0.1343	0.8027	0.1418	0.7972	0.0912	0.5500
T=200						
EM	0.8481	1.4356	0.8422	1.4312	0.8803	1.9568
Bias	-0.0519	-0.5644	-0.0578	-0.5687	-0.0197	-0.0432
RMSE	0.0788	0.6743	0.0824	0.6719	0.0488	0.3871
T=300						
EM	0.8604	1.4857	0.8571	1.4904	0.8832	1.9543
Bias	-0.0396	-0.5143	-0.0430	-0.5096	-0.0168	-0.0457
RMSE	0.0620	0.6030	0.0646	0.6039	0.0391	0.3173

Table 4. Empirical means, bias and root mean squared errors of the estimates of α and θ for PINARPL(1).

Sample size	$\hat{\alpha}_{PCLS}$	$\hat{\theta}_{PCLS}$	$\hat{\alpha}_{PYW}$	$\hat{\theta}_{PYW}$	$\hat{\alpha}_{PCML}$	$\hat{\theta}_{PCML}$
True value $\alpha = 0.2$ and $\theta = 0.6$						
T=100						
EM	0.1725	0.5844	0.1708	0.5896	0.1312	0.6324
Bias	-0.0275	-0.0156	-0.0292	-0.0104	-0.0688	0.0324
RMSE	0.1044	0.0820	0.1041	0.0820	0.0962	0.0774
T=200						
EM	0.1821	0.5880	0.1812	0.5906	0.1271	0.6239
Bias	-0.0179	-0.0119	-0.0188	-0.0094	-0.0729	0.0239
RMSE	0.0756	0.0588	0.0754	0.0586	0.0863	0.0551
T=300						
EM	0.1867	0.5892	0.1860	0.5909	0.1267	0.6227
Bias	-0.0132	-0.0108	-0.0139	-0.0091	-0.0733	0.0227
RMSE	0.0607	0.0491	0.0606	0.0489	0.0819	0.0456
True value $\alpha = 0.5$ and $\theta = 1$						
T=100						
EM	0.4426	0.9084	0.4384	0.9167	0.4424	1.0785
Bias	-0.0573	-0.0915	-0.0616	-0.0832	-0.0576	0.0785
RMSE	0.1175	0.1793	0.1188	0.1768	0.1012	0.1887
T=200						
EM	0.4588	0.9172	0.4563	0.9212	0.4498	1.0761
Bias	-0.0412	-0.0828	-0.0436	-0.0788	-0.0502	0.0761
RMSE	0.0844	0.1395	0.0852	0.1376	0.0783	0.1387
T=300						
EM	0.4595	0.9177	0.4579	0.9203	0.4504	1.0726
Bias	-0.0405	-0.0823	-0.0421	-0.0797	-0.0496	0.0726
RMSE	0.0729	0.1255	0.0738	0.1241	0.0700	0.1193
True value $\alpha = 0.9$ and $\theta = 2$						
T=100						
EM	0.8289	1.4464	0.8178	1.4307	0.6528	2.4360
Bias	-0.0711	-0.5535	-0.0822	-0.5692	-0.2472	0.4360
RMSE	0.1050	0.8810	0.1127	0.6997	0.9387	1.3185
T=200						
EM	0.8572	1.5699	0.8520	1.5733	0.7295	2.4420
Bias	-0.0428	-0.4301	-0.0479	-0.4267	-0.1705	0.4420
RMSE	0.0659	0.5841	0.0695	0.5848	0.2041	0.5434
T=300						
EM	0.8683	1.6127	0.8649	1.6167	0.7427	2.4417
Bias	-0.0317	-0.3873	-0.0350	-0.3830	-0.1573	0.4417
RMSE	0.0505	0.5273	0.0530	0.5322	0.1818	0.5047

2. Due to the Figure 2, one can realize that an INAR(1) process may be appropriate for modeling this data set because exists a cut-off after lag 1 in the sample partial autocorrelation. The sample mean, variance and dispersion index are respectively, 20.02, 52.75 and 2.63. The value of dispersion index shows that the data series is overdispersed. For the earthquakes se-

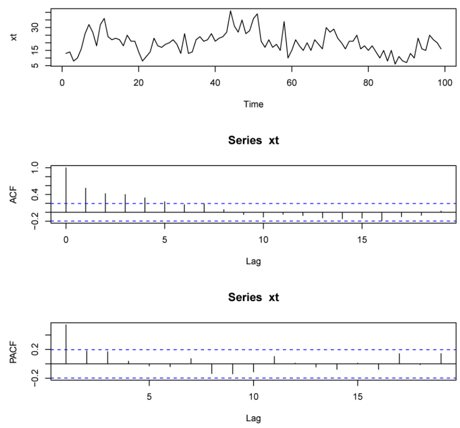

Figure 2. The sample path, ACF and PACF plots of the number of earthquakes magnitude 7.0 or greater, annually from 1900-1998 in the world.

ries under the BINARPL(1), PINARPL(1), NBINARPL(1), INARP(1) and INARG(1) models, we calculate the CML, CLS and YW estimates of the parameters, the Akaike information criterion (AIC) and Bayesian information criterion (BIC). Results are presented in 5. As we can see from 5, the PINARPL(1) model have the smallest the AIC and BIC than the others. Thus, the PINARPL(1) model with $W_t \sim PL(0.2878)$ innovations that gives by

$$X_t = 0.694 \ominus X_{t-1} + W_t,$$

is more appropriate for fit to this data. Where $\ominus$ indicates the Poisson thinning operator. The predicted values of the earthquakes series are given by

$$\begin{aligned}\hat{X}_1 &= \frac{\hat{\theta}+2}{\hat{\theta}(\hat{\theta}+1)(1-\hat{\alpha})} = 20.187,\\ \hat{X}_i &= \hat{\alpha}\hat{X}_{i-1} + \frac{\hat{\theta}+2}{\hat{\theta}(\hat{\theta}+1)} = 0.694\hat{X}_{i-1} + 6.173, \qquad i = 2, 3, ..., 99.\end{aligned}$$

Table 5. Estimated parameters, AIC and BIC for the earthquakes series.

Model	CLS	YW	CML	AIC	BIC
NBINARPL(1)	$\hat{\alpha} = 0.5434$	$\hat{\alpha} = 0.5417$	$\hat{\alpha} = 0.7398$	637.9338	643.1241
	$\hat{\theta} = 0.1969$	$\hat{\theta} = 0.1998$	$\hat{\theta} = 0.3330$		
BINARPL(1)	$\hat{\alpha} = 0.5434$	$\hat{\alpha} = 0.5417$	$\hat{\alpha} = 0.6099$	642.9801	648.1704
	$\hat{\theta} = 0.1969$	$\hat{\theta} = 0.1998$	$\hat{\theta} = 0.2304$		
PINARPL(1)	$\hat{\alpha} = 0.5434$	$\hat{\alpha} = 0.5417$	$\hat{\alpha} = 0.6942$	636.1583	641.3485
	$\hat{\theta} = 0.1969$	$\hat{\theta} = 0.1998$	$\hat{\theta} = 0.2878$		
INARG(1)	$\hat{\alpha} = 0.5434$	$\hat{\alpha} = 0.5417$	$\hat{\alpha} = 0.657$	654.097	659.287
	$\hat{P} = 0.097$	$\hat{P} = 0.098$	$\hat{P} = 0.126$		
INARP(1)	$\hat{\alpha} = 0.5434$	$\hat{\alpha} = 0.5417$	$\hat{\alpha} = 0.3822$	674.5856	679.7758
	$\hat{\lambda} = 9.323$	$\hat{\lambda} = 9.1746$	$\hat{\lambda} = 12.42$		

Figure 3 shows the predicted values to the sample paths of the earthquakes series.

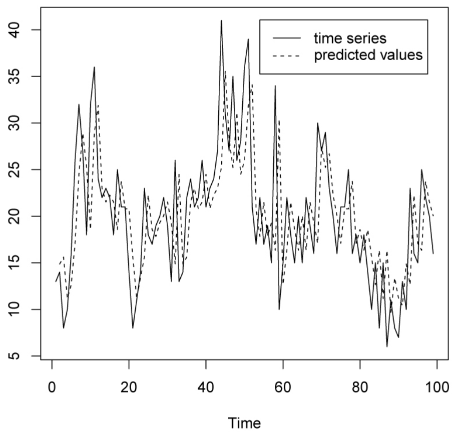

Figure 3. Predicted values of the earthquakes series.

7.2 The Number of Measles Cases

The second example assumes the number of measles cases, monthly from Aug 2013-Dec 2016 in Sweden. The sample path, autocorrelation and partial autocorrelation functions are shown in Figure 4. Due to exists a cut-off after lag 1 in the sample partial autocorrelation in Figure 4, we conclude that an INAR(1) process may be appropriate for modeling the measles series, because of the clear cut-off after lag 1 in the partial autocorrelations. The sample mean, variance and dispersion index are respectively, 1.244, 3.489 and 2.805. The value of dispersion index shows that the data series is overdispersed.

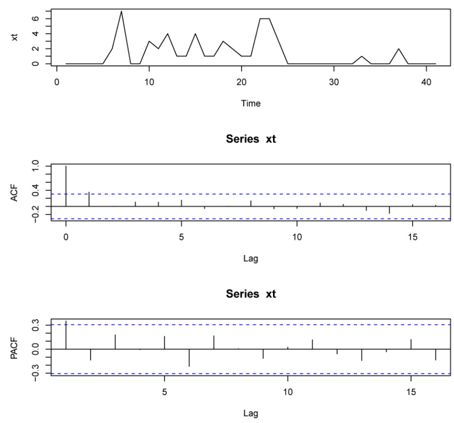

Figure 4. The sample path, ACF and PACF plots of the number of measles cases, monthly from Aug 2013-Dec 2016 in Sweden.

For the BINARPL(1), PINARPL(1), NBINARPL(1), INARP(1) and INARG(1) models, we obtain the CML, CLS and YW estimates of the parameters, the AIC and BIC. The results are shown in Table 6. According to Table 6, since the values of the AIC and BIC are smaller for the NBINARPL(1) model, this model with $W_t \sim PL(2.349)$ innovations that is given by

$$X_t = 0.56 * X_{t-1} + W_t,$$

Table 6. Estimated parameters, AIC and BIC for the measles series.

Model	CLS	YW	MLE	AIC	BIC
NBINARPL(1)	$\hat{\alpha} = 0.355$	$\hat{\alpha} = 0.351$	$\hat{\alpha} = 0.5631$	122.764	126.191
	$\hat{\theta} = 1.671$	$\hat{\theta} = 1.698$	$\hat{\theta} = 2.3490$		
BINARPL(1)	$\hat{\alpha} = 0.355$	$\hat{\alpha} = 0.351$	$\hat{\alpha} = 0.2506$	125.7234	129.1505
	$\hat{\theta} = 1.671$	$\hat{\theta} = 1.698$	$\hat{\theta} = 1.4874$		
PINARPL(1)	$\hat{\alpha} = 0.355$	$\hat{\alpha} = 0.351$	$\hat{\alpha} = 0.3307$	124.8738	128.3010
	$\hat{\theta} = 1.671$	$\hat{\theta} = 1.698$	$\hat{\theta} = 1.6368$		
INARG(1)	$\hat{\alpha} = 0.355$	$\hat{\alpha} = 0.351$	$\hat{\alpha} = 0.248$	124.623	128.050
	$\hat{P} = 0.549$	$\hat{P} = 0.553$	$\hat{P} = 0.510$		
INARP(1)	$\hat{\alpha} = 0.355$	$\hat{\alpha} = 0.351$	$\hat{\alpha} = 0.294$	143.801	147.228
	$\hat{\lambda} = 0.822$	$\hat{\lambda} = 0.807$	$\hat{\lambda} = 0.899$		

being better. The predicted values of the measles series are given by

$$\begin{aligned}\hat{X}_1 &= \frac{\hat{\theta}+2}{\hat{\theta}(\hat{\theta}+1)(1-\hat{\alpha})} = 1.256,\\ \hat{X}_i &= \hat{\alpha}\hat{X}_{i-1} + \frac{\hat{\theta}+2}{\hat{\theta}(\hat{\theta}+1)} = 0.56\hat{X}_{i-1} + 0.55, \qquad i = 2, 3, ..., 41.\end{aligned}$$

Figure 5 shows the predicted values to the sample paths of the measles series.

7.3 The Numbers of Sudden Death Series

This example assumes the numbers of submissions to animal health laboratories, monthly from 2003-2009, from a region in New Zealand. The submissions can be categorized in various ways. Data set is sudden death series and this data is used by Aghababaei Jazi et al. (2012).
The sample path, autocorrelation and partial autocorrelation functions of series are shown in Figure 6. According to Figure 6, an INAR(1) process may be appropriate for modeling this data set because exists a cut-off after lag 1 in the sample partial autocorrelation. The sample mean, variance and dispersion index are respectively, 2.0238, 6.529 and 3.23. The value of dispersion index shows that the data series is overdispersed.

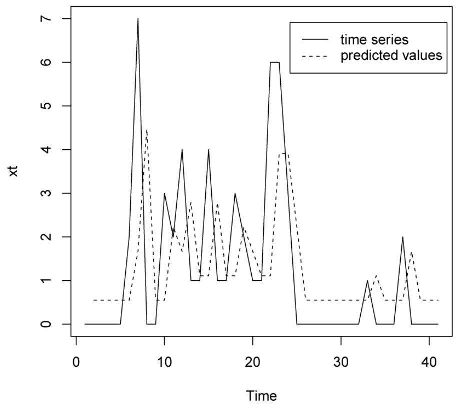

Figure 5. Predicted values of the measles series.

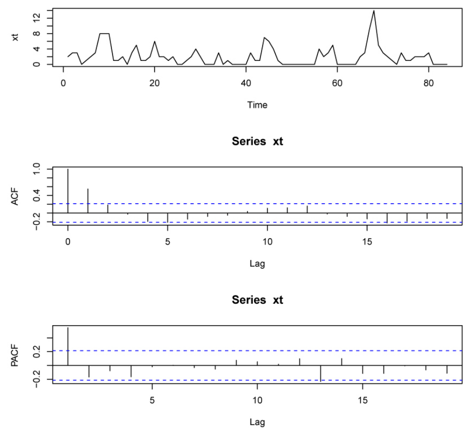

Figure 6. The sample path, ACF and PACF plots of the sudden death series.

in the Table 7, we present the CML, CLS and YW estimates of the parameters, the AIC and BIC for the BINARPL(1), PINARPL(1), NBINARPL(1), INARP(1) and INARG(1) models. According to Table 7, since the values of the AIC and BIC are smaller for the NBINARPL(1) model compared with those values of the other model, this model with $W_t \sim PL(1.6850)$ innovations that is given by

$$X_t = 0.59 * X_{t-1} + W_t,$$

is redan appropriate model for fit to this data series. The predicted values

Table 7. Estimated parameters, AIC and BIC for the sudden death series.

Model	CLS	YW	CML	AIC	BIC
NBINARPL(1)	$\hat{\alpha} = 0.5521$	$\hat{\alpha} = 0.5478$	$\hat{\alpha} = 0.5888$	297.5909	302.4525
	$\hat{\theta} = 1.5221$	$\hat{\theta} = 1.5255$	$\hat{\theta} = 1.6850$		
BINARPL(1)	$\hat{\alpha} = 0.5521$	$\hat{\alpha} = 0.5478$	$\hat{\alpha} = 0.3191$	308.3543	313.2159
	$\hat{\theta} = 1.5221$	$\hat{\theta} = 1.5255$	$\hat{\theta} = 1.0883$		
PINARPL(1)	$\hat{\alpha} = 0.5521$	$\hat{\alpha} = 0.5478$	$\hat{\alpha} = 0.4732$	303.1880	308.0497
	$\hat{\theta} = 1.5221$	$\hat{\theta} = 1.5255$	$\hat{\theta} = 1.3599$		
INARG(1)	$\hat{\alpha} = 0.5521$	$\hat{\alpha} = 0.5478$	$\hat{\alpha} = 0.317$	306.0826	310.9443
	$\hat{P} = 0.5215$	$\hat{P} = 0.5221$	$\hat{P} = 0.421$		
INARP(1)	$\hat{\alpha} = 0.5521$	$\hat{\alpha} = 0.5478$	$\hat{\alpha} = 0.3828$	347.4463	352.308
	$\hat{\lambda} = 0.9174$	$\hat{\lambda} = 0.9151$	$\hat{\lambda} = 1.240$		

of the sudden death series are given by

$$\hat{X}_1 = \frac{\hat{\theta}+2}{\hat{\theta}(\hat{\theta}+1)(1-\hat{\alpha})} = 1.986,$$
$$\hat{X}_i = \hat{\alpha}\hat{X}_{i-1} + \frac{\hat{\theta}+2}{\hat{\theta}(\hat{\theta}+1)} = 0.59\hat{X}_{i-1} + 0.814, \qquad i = 2, 3, ..., 84.$$

Figure 7 shows the predicted values to the sample paths of the sudden death series.

7.4 The Incidence of Acute Febrile Muco-Cutaneous Lymph Node Syndrome Series

The last example assumes weekly counts of the incidence of acute febrile mucocutaneous lymph node syndrome (MCLS) in Totori-prefecture, Japan, during 1982.

The sample path, autocorrelation and partial autocorrelation functions of series are shown in Figure 8. According to Figure 8, Due to exists a cut-off after lag 1 in the sample partial autocorrelation in Figure 8, we conclude that an INAR(1) process may be appropriate for modeling this series, because of

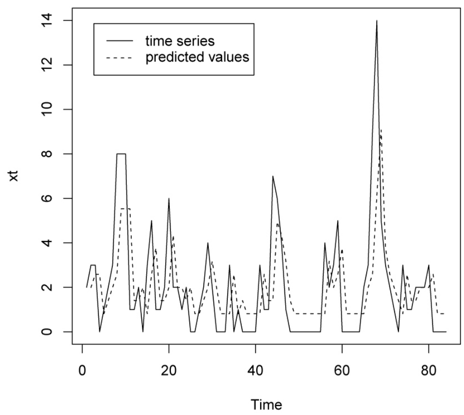

Figure 7. Predicted values of the sudden death series.

the clear cut-off after lag 1 in the partial autocorrelations. The sample mean, variance and dispersion index are respectively, 1.711, 3.111 and 1.818. The value of dispersion index shows that the data series is overdispersed.

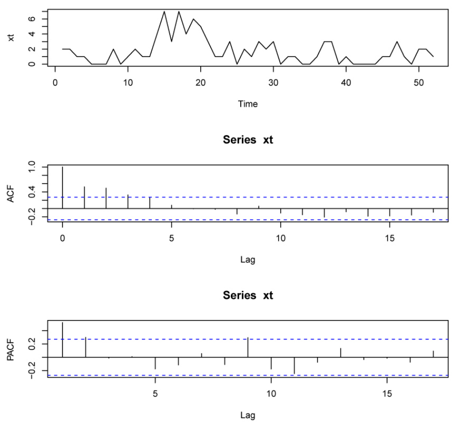

Figure 8. The time series, ACF and PACF plots of the weekly counts of the incidence of acute febrile muco-cutaneous lymph node syndrome (MCLS) in Totori-prefecture, Japan, during 1982.

Table 8 gives the CML, CLS and YW estimates of the parameters, the AIC and BIC for the BINARPL(1), PINARPL(1), NBINARPL(1), INARP(1) and INARG(1) models. According to this table, the NBINARPL(1) model

Table 8. Estimated parameters, AIC and BIC for the incidence of acute febrile MCLS series.

Model	CLS	YW	CML	AIC	BIC
NBINARPL(1)	$\hat{\alpha} = 0.524$	$\hat{\alpha} = 0.522$	$\hat{\alpha} = 0.5209$	170.6369	174.5394
	$\hat{\theta} = 1.640$	$\hat{\theta} = 1.680$	$\hat{\theta} = 1.6908$		
BINARPL(1)	$\hat{\alpha} = 0.524$	$\hat{\alpha} = 0.522$	$\hat{\alpha} = 0.3832$	172.2558	176.1583
	$\hat{\theta} = 1.640$	$\hat{\theta} = 1.680$	$\hat{\theta} = 1.3607$		
PINARPL(1)	$\hat{\alpha} = 0.524$	$\hat{\alpha} = 0.522$	$\hat{\alpha} = 0.4804$	171.0987	175.0012
	$\hat{\theta} = 1.640$	$\hat{\theta} = 1.680$	$\hat{\theta} = 1.5773$		
INARG(1)	$\hat{\alpha} = 0.524$	$\hat{\alpha} = 0.522$	$\hat{\alpha} = 0.3905$	172.5549	176.4574
	$\hat{P} = 0.543$	$\hat{P} = 0.550$	$\hat{P} = 0.492$		
INARP(1)	$\hat{\alpha} = 0.524$	$\hat{\alpha} = 0.522$	$\hat{\alpha} = 0.372$	176.4462	180.3487
	$\hat{\lambda} = 0.841$	$\hat{\lambda} = 0.817$	$\hat{\lambda} = 1.063$		

with $W_t \sim PL(1.691)$ innovations that is given by

$$X_t = 0.521 * X_{t-1} + W_t,$$

is more appropriate for fit to this data because the smallest values of the AIC and BIC are obtained for this model. The predicted values of the incidence of acute febrile MCLS series are given by

$$\begin{aligned} \hat{X}_1 &= \frac{\hat{\theta}+2}{\hat{\theta}(\hat{\theta}+1)(1-\hat{\alpha})} = 1.693, \\ \hat{X}_i &= \hat{\alpha}\hat{X}_{i-1} + \frac{\hat{\theta}+2}{\hat{\theta}(\hat{\theta}+1)} = 0.521\hat{X}_{i-1} + 0.811, \qquad i = 2, 3, ..., 52. \end{aligned}$$

Figure 9 shows the predicted values to the sample paths of the incidence of acute febrile MCLS series.

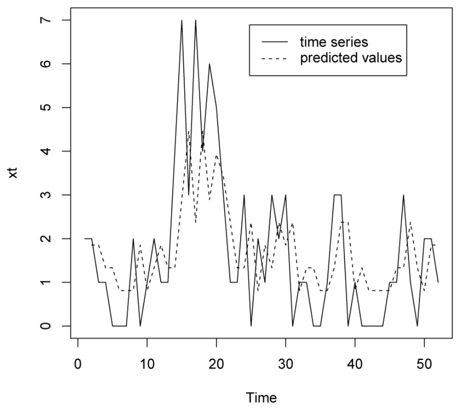

Figure 9. Predicted values of the incidence of acute febrile MCLS series.

8 Conclusion

Integer-valued time series models are very applicable in many fields such as medicine, reliability theory, precipitation, transportation, hotel accommodation and queuing theory. So far, many integer-valued autoregressive processes have been introduced by researchers.

In this paper, we introduce a new stationary first-order integer-valued AR(1) process with Poisson-Lindley innovations based on the power series thinning operator. The main properties of the proposed model are derived. The parameters of the model are estimated using YW, CLS and CML methods. Performance of the estimators are evaluated via simulation. Three special cases of the proposed model (INARPL(1) based on the binomial operator, INARPL(1) based on the Poisson operator and INARPL(1) based on the negative binomial operator) are studied in some detail. Finally, we fitted some submodels of the PSINARPL(1) model to four real data sets to show the potentially of the new proposed model.

Acknowledgment

The authors thank the anonymous referees for their valuable comments and careful reading, which led to an improvement of the presentation and results of the article. The authors are also indebted to Yazd University for supporting this research.

References

Aghababaei Jazi, M.A., Jones, G., and Lai, C.D. (2012). Integer-Valued AR(1) with Geometric Innovations. *Journal of the Iran Statistical Society*, **11**, 173-190.

Al-Osh, M.A., and Alzaid, A.A. (1987). First-Order Integer-Valued Autoregressive (INAR(1)) Process. *Journal of Time Series Analysis*, **8**, 261-275.

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**, 71–93.

Du, J., and Li, Y. (1991). The Integer Valued Autoregressive (INAR(p)) Model. *Journal of Time Series Analysis*, **12**, 129-142.

Ferland, R., Latour, A., and Oraichi, D. (2006). Integer-Valued GARCH Processes. *Journal of Time Series Analysis*, **27**, 923-942.

Lívio, T., Mamode Khan, N., Bourguignon, M., and Bakouch, H.S. (2018). An INAR(1) Model with Poisson-Lindley Innovations. *Economics Bulletin*, **38**, 1505-1513.

Mohammadpour, M., Bakouch, H.S., and Shirozhan, M. (2018). Poisson-Lindley INAR(1) Model with Applications. *Brazilian Journal of Probability and Statistics*, **32**, 262-280.

Ristić, M.M., Bakouch, H.S., and Nastić, A.S. (2009). A New Geometric First Order Integer-Valued Autoregressive (NGINAR(1)) Process. *Statistical Planning and Inference*, **139**, 2218-2226.

Sankaran, M. (1970). *The Discrete Poisson-Lindley Distribution*, vol. 26. Biometrics, Washington, pp. 145-149.

Schweer, S., and Weiß, C.H. (2014). Compound Poisson INAR(1) Processes: Stochastic Properties and Testing for Over Dispersion. *Computational Statistics and Data Analysis*, **77**, 267-284.

Steutel, F.W., and van Harn, K. (1979). Discrete Analogues of Self Decomposability and Stability. *The Annals of Probability*, **7**, 893-899.

Eisa Mahmoudi
Department of Statistics,
Yazd University,
Yazd, Iran.
email: *emahmoudi@yazd.ac.ir*

Ameneh Rostami
Department of Statistics,
Yazd University,
Yazd, Iran.
email: *am64ros@gmail.com*

R. Roozegar
Department of Statistics,
Yazd University,
Yazd, Iran.
email: *r.roozegar@yazd.ac.ir*