Journal of Statistical Theory and Applications
Vol. **19(3)**, September 2020, *pp*. 415–431
DOI: https://doi.org/10.2991/jsta.d.200917.001; ISSN 1538-7887
https://www.atlantis-press.com/journals/jsta

First-Order Integer-Valued Moving Average Process with Power Series Innovations

Eisa Mahmoudi*, Ameneh Rostami

Department of Statistics, Yazd University, P.O. Box 89175-741, Yazd, Iran

ARTICLE INFO

Article History

Received 05 Sep 2019
Accepted 19 May 2020

Keywords

First-order integer-valued moving average process
Poisson thinning operator
Moving average model
Power series family of distributions

2000 Mathematics Subject Classification: 60E05, 62F10.

ABSTRACT

In this paper, we introduce a first-order nonnegative integer-valued moving average process with power series innovations based on a Poisson thinning operator (PINMAPS(1)) for modeling overdispersed, equidispersed and underdispersed count time series. This process contains the PINMA process with geometric, Bernoulli, Poisson, binomial, negative binomial and logarithmic innovations which some of them are studied in details. Some statistical properties of the process are obtained. The unknown parameters of the model are estimated using the Yule-Walker, conditional least squares and least squares feasible generalized methods. Also, the performance of estimators is evaluated using a simulation study. Finally, we apply the model to three real data set and show the ability of the model for predicting data compared to competing models.

© 2020 *The Authors*. Published by Atlantis Press B.V.
This is an open access article distributed under the CC BY-NC 4.0 license (http://creativecommons.org/licenses/by-nc/4.0/).

1. INTRODUCTION

In recent decades, the integer-valued time series have been played an important role in research. This type of time series are extensively used in different science such as natural, social, agricultural, medical and health science. For example, the number of chromosome interchanges in cells, the number of bases of DNA sequences, the number of births in a hospital in successive months, the number of customers in an internet server in a period of time and the number of shares sold in the stock market.

Many models have been proposed by researchers for modeling the integer-valued time series, such as the first-order integer-valued autoregressive (INAR(1)) process (Al-Osh and Alzaid [1]), integer-valued moving average process (INMA(1)) (McKenzie [2]) and (Al-Osh and Alzaid [3]), estimation in INMA models (Brännäs and Hall [4]), bivariate time series modeling of financial count data (Quoreshi [5]), a new geometric INAR(1)) process based on the negative binomial thinning operator (Ristíc *et al.* [6]), integer-valued moving average modeling of the number of transactions in stocks (Brännäs and Quoreshi [7]), acombined geometric INAR(p) model based on the negative binomial thinning (Nastíc *et al.* [8]), a bivariate INAR(1) time series model with geometric marginals (Ristíc *et al.* [9]), compound Poisson INAR(1) processes (Schweer and Wei [10]), INMA models with structural changes (Yu *et al.* [11]), INAR(1) processes with power series innovations (Bourguinon and Vasconcellos [12]), the combined Poisson INMA(q) models for time series of counts (Yu and Zou [13]) and INAR(1) model with Poisson–Lindley marginal distribution (Mohammadpour *et al.* [14]).

One important characteristic of the count data time series is the overdispersion, equidispersion and underdispersion property. The INAR models, are the most widely used integer-valued time series for dealing with this type of data. But INAR models, sometimes, are no more the best choice when data have a very short-run autocorrelation. In this case, the INMA models work better, because in general case for INMA(q), the correlation between Y_t and Y_{t-k} for $k > q$ is zero, whereas, in the INAR models, this correlation gradually decreases with k increase, thus we need to introduce a new model based on the INMA for modeling overdispersed, equidispersed and underdispersed count data that have a very short-run autocorrelation.

The aim of this paper is to introduce a new INMA(1) process with power series innovations based on a Poisson thinning operator. In this process, we use the Poisson thinning operator, which contains Poisson counting series. Unlike the binomial thinning operator (Steutel and van Harn [15]) which counting series can take only 0 or 1 value and is appropriate for modeling the number of random events which could only survive or disappear after a time period, the counting series of the Poisson thinning operator like the negative binomial thinning operator (Ristíc *et al.* [6]) can take any nonnegative integer values and is appropriate for modeling the number of random events capable

*Corresponding author. Email: emahmoudi@yazd.ac.ir

of replication themselves. The main reason for using the Poisson thinning operator instead of the negative binomial thinning operator in this model is that INMA(1) process based on the Poisson thinning operator produced more dependence between the time series variables in comparing with INMA(1) process based on the negative binomial thinning operator under same assumptions.

To clear up the reason why the proposed model in this paper has been used, paying attention to the below points is essential.

i. This model is suitable for modeling time series count data with overdispersion, equidispersion and underdispersion that have a very short-run autocorrelation.

ii. In this model, we use of innovations that come from the power series family of distributions. This family of distributions has the functional form of parameters. Also, it is useful for modeling overdispersed, equidispersed and underdispersed count data. Moreover, the power series family of distributions includes many of the important discrete distributions such as Poisson, geometric, Bernoulli, binomial, negative binomial and logarithmic that are flexible and extensively used distributions for fitting count data.

The paper is organized as follows:

The INMA(1) model with power series innovations based on the Poisson thinning operator is defined in Section 2 and some of its statistical properties are presented. In Section 3, the estimators of the model parameters are obtained using the Yule-Walker (YW), conditional least squares (CLS) and feasible generalized least squares (FGLS) methods. In Section 4, four special cases of the model are given. Some simulation results of the estimators are provided in Section 5. Section 6 deals with three real applications of the proposed model. Finally, we conclude the paper in Section 7.

2. CONSTRUCTION OF THE MODEL

In this section, we introduce an INMA model of the first order generated by the Poisson thinning operator and with the power series innovations. Therefore, we first introduce the Poisson thinning operator.

The Poisson thinning operator, denoted by $\ominus$, is defined by

$$\alpha \ominus X = \sum_{i=1}^{X} Z_i,$$

where $\{Z_i\}_{i=1}^{X}$ is a sequence of independent and identically distributed (i.i.d.) Poisson random variables with mean α and $Z_0 \overset{a.s.}{=} 0$. For a given random variable X, the random variable $\alpha \ominus X$ has the Poisson distribution with mean αX. This implies that the conditional mean and variance of the random variable $\alpha \ominus X$ given X are given as $E(\alpha \ominus X|X) = \alpha X$ and $Var(\alpha \ominus X|X) = \alpha X$, respectively. Then, the unconditional mean and variance of the random variable $\alpha \ominus X$ are, respectively, given by $E(\alpha \ominus X) = \alpha E(X)$ and $Var(\alpha \ominus X) = \alpha E(X) + \alpha^2 Var(X)$. Since the counting series are Poisson distributed random variables, the probability generating function (pgf) of the random variable $\alpha \ominus X$ is given by $\varphi(s) = \varphi_X(e^{\alpha(s-1)})$.

As we mention above, we consider a model which innovations have distributed according to the power series family of distributions, so we review some properties of this family of distributions. A nonnegative integer-valued random variable X is said to have a power series family of distributions, if its probability mass function is given by

$$P(X = x) = \frac{a(x)\theta^x}{C(\theta)}, \; x \in T, \tag{1}$$

where T is a subset of the nonnegative integers, $a(x) \geq 0, \theta > 0$ and $C(\theta)$ is a function defined as $C(\theta) = \sum_{x \in T} a(x)\theta^x$.

This family of distributions contains some well-known distributions such are Bernoulli, binomial, Poisson, geometric, negative binomial and logarithmic distribution. This is shown in Table 1.

Table 1 Special cases of power series distribution.

Distribution	θ	$C(\theta)$	$a(x)$	T
Bernoulli with parameter p	$p/(1-p)$	$1+\theta$	1	$\{0, 1\}$
Binomial with parameters n, p	$p/(1-p)$	$(1+\theta)^n$	$\binom{n}{x}$	$\{0, 1, \dots, n\}$
Poisson with parameter λ	λ	e^{θ}	$1/x!$	$\{0, 1, \dots\}$
Geometric with parameter p	$1-p$	$(1-\theta)^{-1}$	1	$\{0, 1, \dots\}$
Negative binomial with parameter r, p	$1-p$	$(1-\theta)^{-r}$	$\binom{x+r-1}{r-1}$	$\{0, 1, \dots\}$
Logarithmic with parameter p	p	$-log(1-\theta)$	x^{-1}	$\{1, 2, \dots\}$

The pgf of a random variable X with power series family of distributions is given by $\varphi_X(s) = \frac{C(\theta s)}{C(\theta)}$. Using this fact, the expectation and variance of the random variable X are given by $E(X) = \theta G'(\theta)$ and $Var(X) = \theta^2 G''(\theta) + \theta G'(\theta)$, respectively, where $G(\theta) = \log C(\theta)$ and G' and G'' are the first two derivatives of function $G(\theta)$. The dispersion index is given by

$$I_x = \frac{\theta^2 G''(\theta) + \theta G'(\theta)}{\theta G'(\theta)}.$$

Thus, this distribution is overdispersed if $C(\theta) = (1-\theta)^{-1}$, $C(\theta) = (1-\theta)^{-r}$ and $C(\theta) = -log(1-\theta)$ where $-log(1-\theta) > 1$, underdispersed if $C(\theta) = 1+\theta$, $C(\theta) = (1+\theta)^n$ and $C(\theta) = -log(1-\theta)$ where $0 < -log(1-\theta) < 1$, and equidispersed if $C(\theta) = e^{\theta}$.

Definition 2.1. A time series model $\{Y_t\}$ given by

$$Y_t = \alpha \ominus \varepsilon_{t-1} + \varepsilon_t,\ t \in \{0, \pm 1, \pm 2, \dots\}, \tag{2}$$

is called the integer-valued moving average model with power series innovations based on the Poisson thinning operator (*PINMAPS*(1)) if the following conditions are satisfied:

i. $\{\varepsilon_t\}$ is a sequence of i.i.d. random variables with a power series distribution given by (1).

ii. The counting series $\{Z_i^{(t)}\}$ incorporated in $\alpha \ominus \varepsilon_t$ have the Poisson distribution with the parameter $\alpha \in (0,1)$ for all t and i.

iii. All the counting series incorporated in $\alpha \ominus \varepsilon_s$ and $\alpha \ominus \varepsilon_t$ are independent for all $s \neq t$.

iv. The counting series $\{Z_i^{(s)}\}$ are independent of the random variables ε_t for all t, s and i.

Under the above assumptions, we obtain the expectation and variance of the random variable Y_t, respectively, as follows:

$$\begin{aligned} E(Y_t) &= (1+\alpha)\theta G'(\theta), \\ Var(Y_t) &= \alpha\theta G'(\theta) + (1+\alpha^2)[\theta^2 G''(\theta) + \theta G'(\theta)]. \end{aligned}$$

Thus, the dispersion index is given by $I_{Y_t} = 1 + \frac{\theta \frac{G''(\theta)}{G'(\theta)} + \alpha^2[\theta \frac{G''(\theta)}{G'(\theta)} + 1]}{1+\alpha}$.

Remark 2.1

i. If $C(\theta) = (1-\theta)^{-1}$, $C(\theta) = (1-\theta)^{-r}$ and $C(\theta) = e^{\theta}$, the first-order nonnegative integer-valued moving average process with power series innovations based on a Poisson thinning operator (PINMAPS(1)) process is overdispersed.

ii. If $C(\theta) = 1+\theta$ and $C(\theta) = (1+\theta)^n$, this process is overdispersed when $\theta < \alpha^2$; underdispersed when $\theta > \alpha^2$ and equidispersed when $\theta = \alpha^2$.

iii. If $C(\theta) = -log(1-\theta)$, this process is overdispersed when $-log(1-\theta) > 1$ and underdispersed when $0 < -log(1-\theta) < 1$.

The covariance between the random variables Y_t and Y_{t-1} is $Cov(Y_t, Y_{t-1}) = \alpha[\theta^2 G''(\theta) + \theta G'(\theta)]$, which implies that the lag 1 serial correlation of PINMAPS(1) model is

$$\rho(1) = \frac{\alpha\sigma_\varepsilon^2}{\alpha\mu_\varepsilon + (1+\alpha^2)\sigma_\varepsilon^2} = \frac{\alpha[\theta^2 G''(\theta) + \theta G'(\theta)]}{\alpha\theta G'(\theta) + (1+\alpha^2)[\theta^2 G''(\theta) + \theta G'(\theta)]},$$

where μ_ε and σ_ε^2 are the expectation and variance of the innovation ε_t. All other lag $k \geq 2$ serial correlations are equal to 0. We can see that $\rho(1)$ is nonnegative value and bounded above by 1/2, which is the same as for the INMA(1) model introduced by Al-Osh and Alzaid [3].

Theorem 2.1. *Let Y_t be the process defined in (2), then Y_t is covariance stationary.*

Proof. Since the expectation and variance of the process are constant and autocovariance function does not depend on time, the process (2) is covariance stationary.

Theorem 2.2. *The PINMAPS(1) process is ergodic in the mean and autocovariance function.*

Proof. The proof is similar to the proof of theorem 7 from Yu and Zou [13] and omitted.

Using the pgf of the power series distribution and the independency of the counting series and the random variables ε_{t-1} and ε_t, we obtain that the pgf of the random variable Y_t is given by

$$\varphi_{Y_t}(s) = E(s^{\alpha \ominus \varepsilon_{t-1}})E(s^{\varepsilon_t}) = \varphi_{\varepsilon_t}(e^{\alpha(s-1)})\varphi_{\varepsilon_t}(s) = \frac{C(e^{\alpha(s-1)}\theta)}{C(\theta)} \times \frac{C(s\theta)}{C(\theta)}.$$

Using the series expansion of the function C, we obtain

$$C(e^{\alpha(s-1)}\theta) = \frac{a(0)}{C(\theta)} + \sum_{x=1}^{T} \frac{a(x)\theta^x}{C(\theta)} e^{\alpha x(s-1)},$$

which implies that the random variable Y_t is distributed as a random variable $X + S_{\alpha W}$, where the random variables X and $S_{\alpha W}$ are independent random variables, X and W have the power series distributions given by (1) and the random variable $S_{\alpha W}$ for given $W \geq 1$ has Poisson distribution with the mean parameter αW and $S_0 \overset{a.s.}{=} 0$. Then, from the last result, we can easily obtain the probability mass function of the random variable y_t as

$$P(Y_t = y) = \sum_{x=0}^{\min(y,T)} \frac{a(x)\theta^x}{C(\theta)} \left[\frac{a(0)}{C(\theta)} I_{\{x=y\}} + \sum_{w=1}^{T} \frac{a(w)(\alpha w)^{y-x}\theta^w e^{-\alpha w}}{C(\theta)(y-x)!} \right], \quad y \geq 0.$$

Unlike the power series innovations which can take values on set $\{0, 1, \dots, T\}$, the random variable Y_t always takes values on the set of nonnegative integers. This is a consequence of using the Poisson thinning operator.

Al-Osh and Alzaid [3] have shown that the Poisson INMA(1) process is the only INMA(1) process that has a linear regression. An example of the INMA(1) process with a nonlinear regression is the geometric INMA(1) process introduced by Alzaid and Al-Osh [16]. Now, we will consider the regression of the proposed model. In this sense, we first derive the joint pgf of the random variables Y_{t-1} and Y_t. It is given by

$$\varphi_{Y_{t-1},Y_t}(s_1, s_2) = \frac{C(\theta s_2)}{C(\theta)} \times \frac{C(\theta e^{\alpha(s_1-1)})}{C(\theta)} \times \frac{C(\theta s_1 e^{\alpha(s_2-1)})}{C(\theta)}.$$

The joint pgf can be used for derivation of the conditional pgf of the random variable Y_t given Y_{t-1}, which the expression is given in the following theorem.

Theorem 2.3. *The conditional pgf of the random variable Y_t given $Y_{t-1} = x$, $x \in \{0, 1, 2, \dots\}$, is given as*

$$\varphi_{Y_t|Y_{t-1}=x}(s) = \frac{C(\theta s) \sum_{j=0}^{x} \sum_{i=1}^{j} \binom{x}{j} \alpha^j a_{ji} \theta^{i+x-j} e^{-i\alpha+\alpha(x-j)(s-1)} C^{(i)}(\theta e^{-\alpha}) C^{(x-j)}(0)}{C(\theta) \sum_{j=0}^{x} \sum_{i=1}^{j} \binom{x}{j} \alpha^j a_{ji} \theta^{i+x-j} e^{-i\alpha} C^{(i)}(\theta e^{-\alpha}) C^{(x-j)}(0)},$$

where the coefficients a_{ji} are given recurrently as $a_{ji} = i a_{j-1,i} + a_{j-1,i-1}$, $i \in \{2, 3, \dots, j-1\}$ and $a_{j1} = a_{jj} = 1$.

Proof. According to Theorem 1.3.1 from Kocherlakota and Kocherlakota [17], the conditional pgf of the random variable Y_t given $Y_{t-1} = x$ can be derived from the joint pgf of the random variables Y_{t-1} and Y_t as

$$\varphi_{Y_t|Y_{t-1}=x}(s) = \frac{\frac{\partial^x \varphi_{Y_{t-1},Y_t}(0,s)}{\partial s_1^x}}{\frac{\partial^x \varphi_{Y_{t-1},Y_t}(0,1)}{\partial s_1^x}}. \tag{3}$$

Let us first consider the partial derivative $\frac{\partial^x \varphi_{Y_{t-1},Y_t}(s_1,s_2)}{\partial s_1^x}$. By the Leibniz's rule, we have

$$\frac{\partial^x \varphi_{Y_{t-1},Y_t}(s_1, s_2)}{\partial s_1^x} = \frac{C(\theta s_2)}{C^3(\theta)} \sum_{j=0}^{x} \binom{x}{j} \frac{\partial^j C(\theta e^{\alpha(s_1-1)})}{\partial s_1^j} \frac{\partial^{x-j} C(\theta s_1 e^{\alpha(s_2-1)})}{\partial s_1^{x-j}}. \tag{4}$$

It is easy to derive the partial derivatives of the function $C(\theta s_1 e^{\alpha(s_2-1)})$ as

$$\frac{\partial^k C(\theta s_1 e^{\alpha(s_2-1)})}{\partial s_1^k} = \theta^k e^{k\alpha(s_2-1)} C^{(k)}(\theta s_1 e^{\alpha(s_2-1)}), \qquad k = 0, 1, \cdots, x. \tag{5}$$

On the other hand and after some calculations, we obtain that the partial derivatives of the function $C(\theta e^{\alpha(s_1-1)})$ are given by

$$\frac{\partial^k C(\theta e^{\alpha(s_1-1)})}{\partial s_1^k} = \alpha^k \sum_{i=1}^{k} a_{ki} \theta^i e^{i\alpha(s_1-1)} C^{(k)}(\theta e^{\alpha(s_1-1)}), \qquad k = 0, 1, \cdots, x, \tag{6}$$

where the coefficients a_{ki} are given in the statement of theorem. Finally, replacing (5) and (6) in (4) with $s_1 = 0$ and $s_2 = s$, we obtain the numerator of (3). In a similar way, we obtain the denominator of (3) which proves the theorem.

Now, we are able to derive regression of our introduced model. It is given as follows:

Corollary 2.1. *The regression of Y_t given $Y_{t-1} = x$ is a nonlinear function given by*

$$E(Y_t|Y_{t-1} = x) = \frac{\sum_{j=0}^{x}\sum_{i=1}^{j}\binom{x}{j}\alpha^j a_{ji}\theta^{i+x-j}e^{-i\alpha}C^{(i)}(\theta e^{-\alpha})C^{(x-j)}(0)\left[\theta C'(\theta)+\alpha(x-j)C(\theta)\right]}{C(\theta)\sum_{j=0}^{x}\sum_{i=1}^{j}\binom{x}{j}\alpha^j a_{ji}\theta^{i+x-j}e^{-i\alpha}C^{(i)}(\theta e^{-\alpha})C^{(x-j)}(0)}.$$

Proof. The proof follows from Corollary 1.3.1 (Kocherlakota and Kocherlakota [17]) and the previous theorem.

The conditional mean and conditional variance of Y_t given $\mathcal{F}_{t-1}$ are obtained as follows:

$$E(Y_t|\mathcal{F}_{t-1}) = \alpha\varepsilon_{t-1} + \mu_\varepsilon, \tag{7}$$

$$Var(Y_t|\mathcal{F}_{t-1}) = \alpha\varepsilon_{t-1} + \sigma_\varepsilon^2, \tag{8}$$

where $\mathcal{F}_{t-1}$ is the information set at time $t-1$.

3. ESTIMATION OF THE UNKNOWN PARAMETERS

In this section, we derive the estimators of the unknown parameters α and θ using the YW, CLS and FGLS methods. Let $Y_1, \cdots, Y_T$, $T \in N$ be a random sample of size T from the PINMAPS(1) process.

3.1. The YW Method

The YW estimators of the parameters α and θ are obtained via solving the following equations:

$$\bar{Y} - (1+\alpha)\theta G'(\theta) = 0,$$

$$S^2 - \alpha\theta G'(\theta) - (1+\alpha^2)(\theta^2 G''(\theta) + \theta G'(\theta)) = 0,$$

$$r_1 S^2 - \alpha(\theta^2 G''(\theta) + \theta G'(\theta)) = 0,$$

where $\bar{Y}$, S^2 and r_1 are the sample mean, variance and autocorrelation function at lag one, respectively.

Theorem 3.1. *The YW estimators of the parameters α and θ are consistent.*

Proof. According to Theorem 2.2, we have $\bar{Y} \xrightarrow{p} \mu_Y$, $S^2 \xrightarrow{p} \sigma_Y^2$ and $r_1 \xrightarrow{p} \rho(1)$ where $\mu_Y := E(Y_t)$, $\sigma_Y^2 := Var(Y_t)$. Thus, using the properties of convergence in probability, the consistency of the YW estimators is resulted.

3.2. The CLS Method

The CLS estimators of the parameters α and θ are given by minimizing the following function with respect to parameter μ_ε

$$S_{CLS} = \sum_{t=2}^{T} e_{1t}^2,$$

where $\mu_\varepsilon := E(\varepsilon_t)$ and $e_{1t} = Y_t - E(Y_t|\mathcal{F}_{t-1})$.

According to the Section 3 from Brännäs and Quoreshi [7], and Equation 7, e_{1t} is given by

$$e_{1t} = \varepsilon_t - \mu_\varepsilon,$$

so, the CLS estimator of μ_ε is given by

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

The CLS estimators of α and θ are obtained by replacing $\hat{\mu}_{\varepsilon(CLS)}$ in equations,

$$\bar{Y} = (1+\hat{\alpha}_{CLS})\hat{\mu}_{\varepsilon(CLS)},$$
$$\hat{\mu}_{\varepsilon(CLS)} = \hat{\theta}_{CLS}G'(\hat{\theta}_{CLS}).$$

3.3. The FGLS Method

This method includes three stages:

Stage 1: Obtaining the CLS estimator of the parameters.

Stage 2: Minimizing the function

$$S = \sum_{t=2}^{T} e_{2t}^2,$$

with respect to parameter σ_ε^2, where $\sigma_\varepsilon^2 := Var(\varepsilon_t)$ and $e_{2t} = (Y_t - E(Y_t|\mathcal{F}_{t-1}))^2 - Var(Y_t|\mathcal{F}_{t-1})$.

Thus, we have

$$\hat{\sigma}_\varepsilon^2 = \frac{\sum_{t=2}^{T}((\varepsilon_t - \hat{\mu}_{\varepsilon(CLS)})^2 - \hat{\alpha}_{CLS}\varepsilon_{t-1})}{(T-2)}.$$

Stage 3: Minimizing the function

$$S_{FGLS} = \sum_{t=2}^{T} \frac{e_{1t}^2}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})},$$

with respect to parameter μ_ε and obtaining FGLS estimators of α and θ. By minimizing the function S_{FGLS} with respect to μ_ε, the FGLS estimator of μ_ε is given by

$$\hat{\mu}_{\varepsilon(FGLS)} = \frac{\sum_{t=2}^{T} \frac{\varepsilon_t}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}{\sum_{t=2}^{T} \frac{1}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}.$$

So, the FGLS estimators of α and θ are given by replacing $\hat{\mu}_{\varepsilon(FGLS)}$ in equations,

$$\bar{Y} = (1 + \hat{\alpha}_{FLGS})\hat{\mu}_{\varepsilon(FGLS)},$$
$$\hat{\mu}_{\varepsilon(FGLS)} = \hat{\theta}_{FGLS} G'(\hat{\theta}_{FGLS}).$$

4. SPECIAL CASES OF THE PINMAPS(1) PROCESS

In this section, we consider four special cases of the PINMAPS(1) process.

4.1. First-Order Integer-Valued Moving Average Process with Geometric Innovations

If $C(\theta) = \frac{1}{1-\theta}(\theta \in (0,1))$, ε_t has geometric distribution, then Y_t is called first-order integer-valued moving average process with geometric innovations (PINMAG(1)) process.

The mean, variance and pgf of this process are given by

$$E(Y_t) = (1+\alpha)\frac{\theta}{1-\theta},$$
$$Var(Y_t) = (\alpha)\frac{\theta}{1-\theta} + (1+\alpha^2)\frac{\theta}{(1-\theta)^2},$$
$$\varphi_{Y_t}(s) = \frac{C(e^{\alpha(s-1)}\theta)}{C(\theta)}\frac{C(s\theta)}{C(\theta)} = \frac{(1-\theta)^2}{(1-s\theta)(1-e^{\alpha(s-1)}\theta)}.$$

The autocovariance function, conditional mean and variance of Y_t are given, respectively, by

$$\gamma_{Y_t}(1) = \alpha\frac{\theta}{(1-\theta)^2},$$
$$E(Y_t|\mathcal{F}_{t-1}) = \alpha\varepsilon_{t-1} + \frac{\theta}{1-\theta},$$
$$Var(Y_t|\mathcal{F}_{t-1}) = \alpha\varepsilon_{t-1} + \frac{\theta}{(1-\theta)^2}.$$

The YW estimators of α and θ are obtained via solving the following equations:

$$\overline{Y} - \frac{(1+\alpha)\theta}{1-\theta} = 0, \tag{9}$$

$$S^2 - \frac{\alpha\theta}{1-\theta} - \frac{(1+\alpha^2)\theta}{(1-\theta)^2} = 0, \tag{10}$$

$$r_1 S^2 - \frac{\alpha\theta}{(1-\theta)^2} = 0. \tag{11}$$

To facilitate obtaining the estimator of θ, the auxiliary parameter $\nu = \frac{\theta}{1-\theta}$ may be used. Therefore, Equations (9–11) can be written as follows:

$$\begin{aligned}
&\overline{Y} - (1+\alpha)\nu = 0,\\
&S^2 - \alpha\nu - (1+\alpha^2)(\nu+\nu^2) = 0,\\
&r_1 S^2 - \alpha(\nu+\nu^2) = 0.
\end{aligned}$$

So, the YW estimators of ν and α are given by

$$\begin{aligned}
\hat{\nu}_{YW} &= \frac{(\overline{Y}-1) + \sqrt{(1-\overline{Y})^2 - 4(r_1 S^2 - \overline{Y})}}{2},\\
\hat{\alpha}_{YW} &= \frac{\overline{Y} - \hat{\nu}_{YW}}{\hat{\nu}_{YW}}.
\end{aligned}$$

The YW estimator for θ is

$$\hat{\theta}_{YW} = \frac{\hat{\nu}_{YW}}{1+\hat{\nu}_{YW}}.$$

According to the previous section, the CLS and FGLS estimators of α and θ are obtained, respectively, by

$$\hat{\theta}_{CLS} = \frac{\frac{\sum_{t=2}^{T}\varepsilon_t}{T-1}}{1+\frac{\sum_{t=2}^{T}\varepsilon_t}{T-1}}, \hat{\alpha}_{CLS} = \frac{\bar{Y} - \frac{\sum_{t=2}^{T}\varepsilon_t}{T-1}}{\frac{\sum_{t=2}^{T}\varepsilon_t}{T-1}},$$

and

$$\hat{\alpha}_{FGLS} = \frac{\bar{Y} - \frac{\sum_{t=2}^{T}\frac{\varepsilon_t}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}{\sum_{t=2}^{T}\frac{1}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}}{\frac{\sum_{t=2}^{T}\frac{\varepsilon_t}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}{\sum_{t=2}^{T}\frac{1}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}}, \hat{\theta}_{FGLS} = \frac{\frac{\sum_{t=2}^{T}\frac{\varepsilon_t}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}{\sum_{t=2}^{T}\frac{1}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}}{1+\frac{\sum_{t=2}^{T}\frac{\varepsilon_t}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}{\sum_{t=2}^{T}\frac{1}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}},$$

4.2. First-Order Integer-Valued Moving Average Process with Poisson Innovations

If $C(\theta) = e^{\theta}(\theta > 0)$, ε_t has Poisson distribution, then Y_t is called first-order integer-valued moving average process with Poisson innovations (PINMAP(1)) process.

The mean, variance and pgf of this process are given by

$$\begin{aligned}
&E(Y_t) = (1+\alpha)\theta,\\
&Var(Y_t) = \alpha\theta + (1+\alpha^2)\theta,\\
&\varphi_{Y_t}(s) = \frac{C(e^{\alpha(s-1)}\theta)}{C(\theta)}\frac{C(s\theta)}{C(\theta)} = \frac{e^{e^{\alpha(s-1)}\theta}e^{s\theta}}{e^{2\theta}}.
\end{aligned}$$

The autocovariance function, conditional mean and variance of Y_t are given, respectively, by

$$\gamma_{Y_t}(1) = \alpha\theta,$$
$$E(Y_t|\mathcal{F}_{t-1}) = \alpha\varepsilon_{t-1} + \theta,$$
$$Var(Y_t|\mathcal{F}_{t-1}) = \alpha\varepsilon_{t-1} + \theta.$$

Since in this process the conditional mean is the same as the conditiona variance, it can be concluded that Y_t given $\mathcal{F}_{t-1}$ has Poisson distribution with mean $\alpha\varepsilon_{t-1} + \theta$.

The YW estimators of α and θ are obtained via solving the following equations:

$$\bar{Y} - (1+\alpha)\theta = 0,$$
$$S^2 - \alpha\theta - (1+\alpha^2)\theta = 0,$$
$$r_1 S^2 - \alpha\theta = 0.$$

Thus, the YW estimators of θ and α are

$$\hat{\theta}_{YW} = \bar{Y} - r_1 S^2,$$
$$\hat{\alpha}_{YW} = \frac{r_1 S^2}{\hat{\theta}_{YW}}.$$

According to the previous section, the CLS and FGLS estimators of α and θ are obtained, respectively, by

$$\hat{\alpha}_{CLS} = \frac{\bar{Y} - \frac{\sum_{t=2}^{T}\varepsilon_t}{T-1}}{\frac{\sum_{t=2}^{T}\varepsilon_t}{T-1}}, \ \hat{\theta}_{CLS} = \frac{\sum_{t=2}^{T}\varepsilon_t}{T-1},$$

and

$$\hat{\alpha}_{FGLS} = \frac{\bar{Y} - \frac{\sum_{t=2}^{T}\frac{\varepsilon_t}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}{\sum_{t=2}^{T}\frac{1}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}}{\frac{\sum_{t=2}^{T}\frac{\varepsilon_t}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}{\sum_{t=2}^{T}\frac{1}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}}, \ \hat{\theta}_{FGLS} = \frac{\sum_{t=2}^{T}\frac{\varepsilon_t}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}{\sum_{t=2}^{T}\frac{1}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}},$$

4.3. First-Order Integer-Valued Moving Average Process with Binomial Innovations

If $C(\theta) = (1+\theta)^n (\theta > 0)$, ε_t has binomial distribution, then Y_t is called first-order integer-valued moving average process with binomial innovations (PINMAB(1)) process.

The mean, variance and pgf of this process are given by

$$E(Y_t) = (1+\alpha)\frac{n\theta}{1+\theta},$$
$$Var(Y_t) = \alpha\frac{n\theta}{1+\theta} + (1+\alpha^2)\frac{n\theta}{(1+\theta)^2},$$
$$\varphi_{Y_t}(s) = \frac{C(e^{\alpha(s-1)}\theta)}{C(\theta)}\frac{C(s\theta)}{C(\theta)} = \frac{(1+e^{\alpha(s-1)}\theta)^n(1+s\theta)^n}{(1+\theta)^{2n}}.$$

The autocovariance function, conditional mean and conditional variance of Y_t are given, respectively, by

$$\gamma_{Y_t}(1) = \alpha\frac{n\theta}{(1+\theta)^2},$$
$$E(Y_t|\mathcal{F}_{t-1}) = \alpha\varepsilon_{t-1} + \frac{n\theta}{1+\theta},$$
$$Var(Y_t|\mathcal{F}_{t-1}) = \alpha\varepsilon_{t-1} + \frac{n\theta}{(1+\theta)^2}.$$

Assuming n is known, the estimators of the parameter α and θ are obtained. The YW estimators of α and θ are obtained via solving the following equations:

$$\bar{Y} - \frac{(1+\alpha)n\theta}{1+\theta} = 0, \tag{12}$$

$$S^2 - \frac{\alpha n\theta}{1+\theta} - \frac{(1+\alpha^2)n\theta}{(1+\theta)^2} = 0, \tag{13}$$

$$r_1 S^2 - \frac{\alpha n\theta}{(1+\theta)^2} = 0. \tag{14}$$

To facilitate obtaining the YW estimator of θ, the auxiliary parameter $\tau = \frac{\theta}{1+\theta}$ may be used. Therefore, Equations (12–14) can be written as follows:

$$\overline{Y} - (1+\alpha)n\tau = 0,$$
$$S^2 - \alpha n\tau - n(1+\alpha^2)(\tau - \tau^2) = 0,$$
$$r_1 S^2 - n\alpha(\tau - \tau^2) = 0.$$

So, the YW estimators of τ and α are

$$\hat{\tau}_{YW} = \frac{(n+\overline{Y}) - \sqrt{(n+\overline{Y})^2 - 4n(\overline{Y} - r_1 S^2)}}{2n},$$
$$\hat{\alpha}_{YW} = \frac{\overline{Y} - n\hat{\tau}_{YW}}{n\hat{\tau}_{YW}}.$$

The YW estimator for parameter θ is given by

$$\hat{\theta}_{YW} = \frac{\hat{\tau}_{YW}}{1 - \hat{\tau}_{YW}}.$$

According to the previous section, the CLS and FGLS estimators of α and θ are obtained, respectively, by

$$\hat{\alpha}_{CLS} = \frac{\bar{Y} - \frac{\sum_{t=2}^{T}\varepsilon_t}{T-1}}{\frac{\sum_{t=2}^{T}\varepsilon_t}{T-1}}, \ \hat{\theta}_{CLS} = \frac{\frac{\sum_{t=2}^{T}\varepsilon_t}{T-1}}{n - \frac{\sum_{t=2}^{T}\varepsilon_t}{T-1}},$$

and

$$\hat{\alpha}_{FGLS} = \frac{\bar{Y} - \frac{\sum_{t=2}^{T}\frac{\varepsilon_t}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}{\sum_{t=2}^{T}\frac{1}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}}{\frac{\sum_{t=2}^{T}\frac{\varepsilon_t}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}{\sum_{t=2}^{T}\frac{1}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}}, \ \hat{\theta}_{FGLS} = \frac{\frac{\sum_{t=2}^{T}\frac{\varepsilon_t}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}{\sum_{t=2}^{T}\frac{1}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}}{n - \frac{\sum_{t=2}^{T}\frac{\varepsilon_t}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}{\sum_{t=2}^{T}\frac{1}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}},$$

4.4. First-Order Integer-Valued Moving Average Process with Negative Binomial Innovations

For $C(\theta) = (1-\theta)^{-r}(\theta \in (0,1))$, ε_t has geometric distribution, then Y_t is called first-order integer-valued moving average process with negative binomial innovations (PINMANB(1)) process.

The mean, variance and pgf of this process are given by

$$E(Y_t) = (1+\alpha)\frac{r\theta}{1-\theta},$$
$$Var(Y_t) = (\alpha)\frac{r\theta}{1-\theta} + (1+\alpha^2)\frac{r\theta}{(1-\theta)^2},$$
$$\varphi_{Y_t}(s) = \frac{C(e^{\alpha(s-1)}\theta)}{C(\theta)}\frac{C(s\theta)}{C(\theta)} = \frac{(1-\theta)^{2r}}{(1-s\theta)^r(1-e^{\alpha(s-1)}\theta)^r}.$$

Also, the autocovariance function, conditional mean and variance of Y_t are given, respectively, by

$$\gamma_{Y_t}(1) = \alpha\frac{r\theta}{(1-\theta)^2},$$
$$E(Y_t|\mathcal{F}_{t-1}) = \alpha\varepsilon_{t-1} + \frac{r\theta}{1-\theta},$$
$$Var(Y_t|\mathcal{F}_{t-1}) = \alpha\varepsilon_{t-1} + \frac{r\theta}{(1-\theta)^2}.$$

Assuming r is known, the estimators of the parameters α and θ are obtained. The YW estimators of α and θ are obtained via solving the following equations:

$$\bar{Y} - \frac{(1+\alpha)r\theta}{1-\theta} = 0, \tag{15}$$

$$S^2 - \frac{\alpha r\theta}{1-\theta} - \frac{(1+\alpha^2)r\theta}{(1-\theta)^2} = 0, \tag{16}$$

$$r_1S^2 - \frac{\alpha r\theta}{(1-\theta)^2} = 0. \tag{17}$$

To facilitate obtaining the YW estimator of θ, the auxiliary parameter $\delta = \frac{\theta}{1-\theta}$ may be used. Therefore, Equations (15–17) can be written as follows:

$$\bar{Y} - (1+\alpha)r\delta = 0,$$
$$S^2 - \alpha r\delta - r(1+\alpha^2)(\delta+\delta^2) = 0,$$
$$r_1S^2 - r\alpha(\delta+\delta^2) = 0.$$

So, the YW estimators of δ and α are

$$\hat{\delta}_{YW} = \frac{(\bar{Y}-r) + \sqrt{(r-\bar{Y})^2 - 4r(r_1S^2-\bar{Y})}}{2r},$$
$$\hat{\alpha}_{YW} = \frac{\bar{Y} - r\hat{\delta}_{YW}}{r\hat{\delta}_{YW}}.$$

The YW estimator for parameter θ is given by

$$\hat{\theta}_{YW} = \frac{\hat{\delta}_{YW}}{1+\hat{\delta}_{YW}}.$$

According to the previous section, the CLS and FGLS estimators of α and θ are obtained, respectively, by

$$\hat{\alpha}_{CLS} = \frac{\overline{Y} - \frac{\sum_{t=2}^{T}\varepsilon_t}{T-1}}{\frac{\sum_{t=2}^{T}\varepsilon_t}{T-1}},\ \hat{\theta}_{CLS} = \frac{\frac{\sum_{t=2}^{T}\varepsilon_t}{T-1}}{r + \frac{\sum_{t=2}^{T}\varepsilon_t}{T-1}},$$

and

$$\hat{\alpha}_{FGLS} = \frac{\overline{Y} - \frac{\sum_{t=2}^{T}\frac{\varepsilon_t}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}{\sum_{t=2}^{T}\frac{1}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}}{\frac{\sum_{t=2}^{T}\frac{\varepsilon_t}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}{\sum_{t=2}^{T}\frac{1}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}},\ \hat{\theta}_{FGLS} = \frac{\frac{\sum_{t=2}^{T}\frac{\varepsilon_t}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}{\sum_{t=2}^{T}\frac{1}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}}{r + \frac{\sum_{t=2}^{T}\frac{\varepsilon_t}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}{\sum_{t=2}^{T}\frac{1}{\widehat{Var}(Y_t|\mathcal{F}_{t-1})}}},$$

Table 2 Estimated parameters, Bias and RMS (in parentheses) for the PINMAP(1) model.

	YW		CLS		FGLS	
T	$\hat{\alpha}_{YW}$	$\hat{\theta}_{YW}$	$\hat{\alpha}_{CLS}$	$\hat{\theta}_{CLS}$	$\hat{\alpha}_{FGLS}$	$\hat{\theta}_{FGLS}$
			$\alpha = 0.4, \theta = 1$			
100	0.2058(0.2655)	1.3060(0.3669)	0.5478(0.1681)	1.0005(0.0960)	0.5491(0.1926)	1.0026(0.1087)
Bias	−0.1942	0.3060	0.1478	0.0005	0.1491	0.0026
200	0.2004(0.2350)	1.3007(0.3332)	0.5455(0.1568)	1.0020(0.0697)	0.5460(0.1671)	1.0030(0.0792)
Bias	−0.1996	0.3007	0.1455	0.0020	0.1460	0.0030
300	0.1989(0.2232)	1.3003(0.3219)	0.5474(0.1547)	1.0023(0.0590)	0.5475(0.1621)	1.0032(0.0664)
Bias	−0.2011	0.3003	0.1474	0.0023	0.1475	0.0032
			$\alpha = 0.7, \theta = 3$			
100	0.7173(0.5780)	3.2053(0.7794)	0.7061(0.0497)	3.0069(0.1721)	0.6932(0.3595)	2.9924(0.4516)
Bias	0.0173	0.2053	0.0061	0.0069	−0.0068	−0.0076
200	0.6639(0.3392)	3.1806(0.5663)	0.7096(0.0355)	2.9997(0.1270)	0.7101(0.0563)	3.0009(0.1469)
Bias	−0.0361	0.1806	0.0096	−0.0003	0.0101	0.0009
300	0.6734(0.2483)	3.1216(0.4439)	0.7104(0.0294)	2.9963(0.1027)	0.7124(0.0451)	2.9939(0.1181)
Bias	−0.0265	0.1216	0.0104	−0.0037	0.0124	−0.0061
			$\alpha = 0.8, \theta = 4$			
100	0.8993(1.2971)	4.2544(1.1438)	0.7987(0.0449)	3.9985(0.1972)	0.7934(0.1754)	3.9904(0.8252)
Bias	0.0993	0.2544	−0.0012	−0.0015	−0.0065	−0.0095
200	0.8141(0.4376)	4.1525(0.8417)	0.8002(0.0315)	3.9978(0.1434)	0.7741(0.5837)	3.9803(0.8145)
Bias	0.0141	0.1525	0.0002	−0.0022	−0.0259	−0.0197
300	0.8147(0.3147)	4.0835(0.6595)	0.8006(0.0249)	4.0064(0.1200)	0.8007(0.0599)	4.0122(0.2408)
Bias	0.0147	0.0835	0.0006	0.0064	0.0007	0.0122

Table 3 Estimated parameters, Bias and RMS (in parentheses) for the PINMAB(1) model.

	YW		CLS		FGLS	
T	$\hat{\alpha}_{YW}$	$\hat{\theta}_{YW}$	$\hat{\alpha}_{CLS}$	$\hat{\theta}_{CLS}$	$\hat{\alpha}_{FGLS}$	$\hat{\theta}_{FGLS}$
			$\alpha = 0.3, \theta = 0.5, n = 5$			
100	0.2466(0.1740)	0.5716(0.1371)	0.3222(0.0506)	0.5015(0.0495)	0.3224(0.0601)	0.5020(0.0523)
Bias	−0.0534	0.0716	0.0222	0.0015	0.0224	0.0020
200	0.2269(0.1451)	0.5762(0.1247)	0.3251(0.0410)	0.5001(0.0341)	0.3251(0.0464)	0.5004(0.0364)
Bias	−0.0730	0.0762	0.0251	0.0001	0.0251	0.0004
300	0.2314(0.1226)	0.5683(0.1036)	0.3231(0.0350)	0.5016(0.0286)	0.3232(0.0392)	0.5017(0.0303)
Bias	−0.0686	0.0684	0.0231	0.0016	0.0232	0.0017
			$\alpha = 0.4, \theta = 0.8, n = 5$			
100	0.3604(0.2118)	0.9178(0.2833)	0.4063(0.0422)	0.8005(0.0737)	0.3933(0.5406)	0.8035(0.1715)
Bias	−0.0396	0.1178	0.0063	0.0005	−0.0067	0.0035
200	0.3559(0.1488)	0.8952(0.2148)	0.4099(0.0328)	0.8028(0.0518)	0.4102(0.0611)	0.8052(0.0694)
Bias	−0.0441	0.0953	0.0099	0.0028	0.0102	0.0052
300	0.3563(0.1271)	0.8787(0.1682)	0.4077(0.0265)	0.8006(0.0424)	0.4082(0.0378)	0.8009(0.0495)
Bias	−0.0437	0.0787	0.0077	0.0006	0.0082	0.0009
			$\alpha = 0.7, \theta = 0.9, n = 5$			
100	0.6523(0.3188)	1.1090(0.5563)	0.7081(0.0569)	0.9034(0.0816)	0.6860(1.2838)	0.8839(0.5572)
Bias	−0.0477	0.2089	0.0082	0.0034	−0.0140	−0.0161
200	0.6485(0.2197)	1.0316(0.3248)	0.7092(0.0390)	0.9033(0.0557)	0.7184(0.3054)	0.9151(0.2037)
Bias	−0.0514	0.1315	0.0092	0.0033	0.0184	0.0151
300	0.6426(0.1837)	1.0162(0.2531)	0.7097(0.0332)	0.9009(0.0475)	0.7149(0.0777)	0.9009(0.0827)
Bias	−0.0574	0.1162	0.0097	0.0009	0.0149	0.0009

Note: RMS, root mean square; PINMAP(1), first-order nonnegative integer-valued moving average process with power series innovations based on a Poisson thinning operator; YW, Yule-Walker; CLS, conditional least squares; FGLS, feasible generalized least squares Reply

5. SIMULATION STUDY

In this section, the performance of the YW, CLS and FGLS estimators is evaluated. For this purpose, we simulate 1000 samples of size $T = 100, 200, 300$ from the PINMAP(1), PINMAB(1) and PINMANB(1) models for different values of α and θ and estimate the parameters using the three methods. Tables 2–4 present the estimates, biases and root mean square (RMS) errors of the estimators. As we see from Tables 2–4, in each of the three models, estimates converge to the true value and the bias and RMS of the CLS estimators are smaller than the bias and RMS of the other two methods, thus the CLS estimators present more performance.

Table 4 Estimated parameters, Bias and RMS (in parentheses) for the PINMANB(1) model.

	YW		CLS		FGLS	
T	$\hat{\alpha}_{YW}$	$\hat{\theta}_{YW}$	$\hat{\alpha}_{CLS}$	$\hat{\theta}_{CLS}$	$\hat{\alpha}_{FGLS}$	$\hat{\theta}_{FGLS}$
			$\alpha = 0.2, \theta = 0.3, r = 10$			
100	0.2106(0.1986)	0.3017(0.0328)	0.1992(0.0229)	0.3002(0.0123)	0.1986(0.0234)	0.3003(0.0124)
Bias	0.0106	0.0017	−0.0008	0.0002	−0.0013	0.00034
200	0.1993(0.1259)	0.3016(0.0225)	0.1999(0.01601)	0.2999(0.0084)	0.1996(0.0165)	0.30004(0.0084)
Bias	−0.0007	0.0016	0.0001	−0.00001	−0.0003	0.00004
300	0.1917(0.0949)	0.3023(0.0174)	0.1997(0.0129)	0.3000(0.0071)	0.1996(0.0132)	0.30003(0.0071)
Bias	−0.0083	0.0023	−0.0002	−0.00001	−0.0004	0.00003
			$\alpha = 0.3, \theta = 0.3, r = 10$			
100	0.2911(0.4048)	0.3046(0.0387)	0.3003(0.0277)	0.3004(0.0122)	0.2990(0.0288)	0.3006(0.0123)
Bias	−0.0089	0.0046	0.0003	0.0004	−0.0009	0.0006
200	0.3024(0.1770)	0.3019(0.0253)	0.2999(0.01967)	0.3002(0.0085)	0.2997(0.0207)	0.3002(0.0085)
Bias	0.0024	0.0019	−0.0001	0.0002	−0.0003	0.0002
300	0.3031(0.1367)	0.3006(0.0217)	0.3010(0.0153)	0.2994(0.0068)	0.3010(0.0161)	0.2995(0.0069)
Bias	0.0031	0.0006	0.0010	−0.0005	0.0010	−0.0005
			$\alpha = 0.5, \theta = 0.4, r = 10$			
100	0.5859(0.6263)	0.3997(0.0611)	0.4954(0.0296)	0.4001(0.0122)	0.4936(0.0313)	0.4004(0.0123)
Bias	0.0859	−0.0003	−0.0045	0.0001	−0.0063	0.0003
200	0.5589(0.4651)	0.3972(0.0441)	0.4966(0.0197)	0.4003(0.0086)	0.4959(0.0209)	0.4004(0.0087)
Bias	0.0589	−0.0028	−0.0033	0.0003	−0.0040	0.0004
300	0.5257(0.2669)	0.3996(0.0351)	0.4995(0.0157)	0.4000(0.0070)	0.4992(0.0167)	0.4001(0.0071)
Bias	0.0257	−0.0004	−0.0004	0.00003	−0.0008	0.0001

Note: RMS, root mean square; PINMAP(1), first-order nonnegative integer-valued moving average process with power series innovations based on a Poisson thinning operator; YW, Yule-Walker; CLS, conditional least squares; FGLS, feasible generalized least squares

6. APPLICATION

In this section, we fit PINMAPS(1) model to three real data sets and obtain the model that gives a better fit to count data. For this purposes, we compare PINMAP(1), PINMAG(1), PINMABE(1) (INMA(1) with Bernoulli innovations based on the Poisson thinning operator), PINMA(1) (INMA(1) with Poisson innovations based on the binomial thinning operator), proposed by Al-Osh and Alzaid [3], GINMA(1) (INMA(1) with geometric innovations based on the binomial thinning operator), proposed by Alzaid and Al-Osh [16], NBINMAP(1) (INMA(1) with Poisson innovations based on the negative binomial thinning operator), NBINMAG(1) (INMA(1) with geometric innovations based on the negative binomial thinning operator) and NBINMABE(1) (INMA(1) with Bernoulli innovations based on the negative binomial thinning operator) models.

6.1. Number of Polio Cases

The first example assumes the number of polio cases, monthly from Jan 1980 to Dec 1983 in the United States A (https://books.google.com/books/about/Multivariate_Statistical_Modelling_Based.html). The sample path, autocorrelation and partial autocorrelation functions are shown in Figure 1. According to Figure 1, we observe that an INMA(1) process can be suitable for modeling the polio series since there exists a cut-off after lag 1 in the sample autocorrelation. The sample mean, variance and empirical index of dispersion are, respectively, 0.7708, 1.2868 and 1.6693. Since the index of dispersion exceeds 1, the polio series is overdispersed. Thus, an overdispersed model must be assumed for modeling the series. We fit the PINMAP(1), PINMAG(1), PINMA(1), GINMA(1), NBINMAP(1) and NBINMAG(1) models to this data set. For the INMA(1) models mentioned, we obtain the YW estimates of the unknown parameters (since the CLS and FGLS estimates are dependent on innovations and we do not observe their values) and the RMS of differences of observations and predicted values (RMS). The results are presented in Table 5. According to this table, we observed that the PINMAG(1) model gives the lowest RMS value compared to the other models. Thus, it can be concluded that the PINMAG(1) model presents the best forecasting for the polio series. Figure 2 shows the plots of the polio data series and their predicted values based on the PINMAG(1) model.

6.2. Number of Rubella Cases

The second example assumes the number of rubella cases, monthly from Jan 2013 to Jul 2017 in Spain (https://www.ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/measles-rubella-monitoring-jan-(2013-2018).pdf). The sample path, autocorrelation and partial autocorrelation functions are shown in Figure 3. According to Figure 3, we observe that an INMA(1) process can be appropriate for modeling the rubella series since there exists a cut-off after lag 1 in the sample autocorrelation. The sample mean, variance and empirical index of dispersion are, respectively, 0.291, 0.321 and 1.104. Since the index of dispersion exceeds 1, the rubella series is

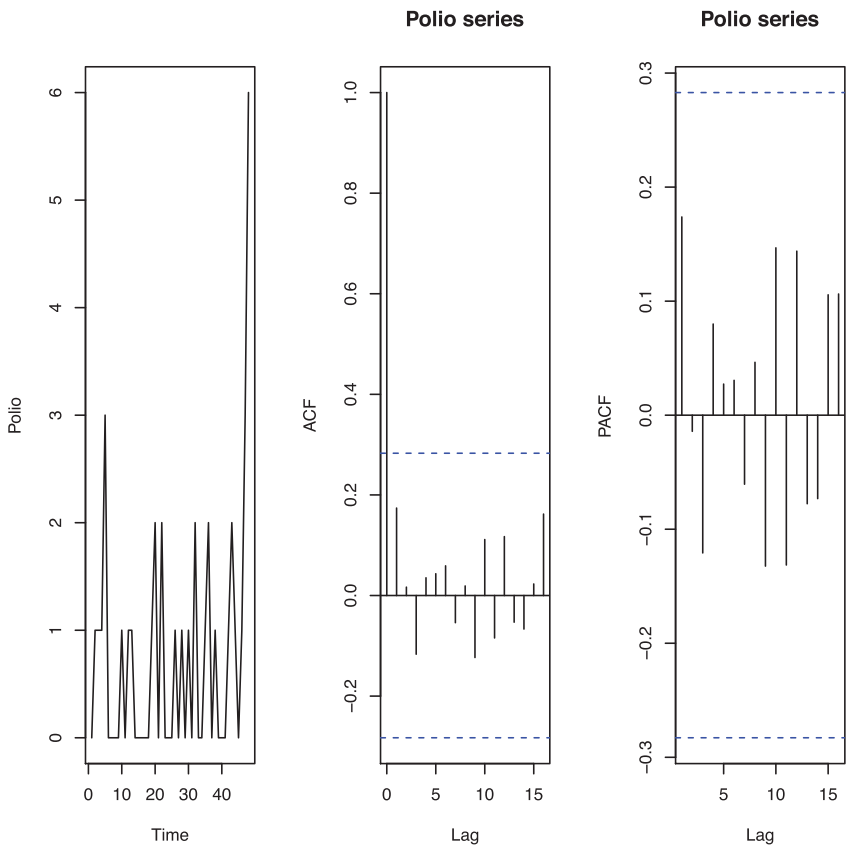

Figure 1 The sample path, Autocorrelation function (ACF) and Partial autocorrelation functions (PACF) plots of the number of polio cases, monthly from Jan 1980 to Dec 1983 in the United States.

Table 5 YW estimates of the parameters and RMS for the polio series.

Model	YW Estimates			RMS
	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{p}$	
PINMAP(1)	0.221	1.014	–	1.156
PINMAG(1)	0.201	–	0.609	1.085
PINMA(1)	0.210	1.063	–	1.175
GINMA(1)	0.196	–	0.608	1.103
NBINMAP(1)	0.466	0.479	–	1.112
NBINMAG(1)	0.207	–	0.610	1.183

RMS, root mean square; PINMAP(1), first-order nonnegative integer-valued moving average process with power series innovations based on a Poisson thinning operator; YW, Yule-Walker

overdispersed. Thus, the series must be modeled by an overdispersed model. We fit the PINMAP(1), PINMAG(1), PINMA(1), GINMA(1), NBINMAP(1) and NBINMAG(1) models to this data set. For each model, we obtain the YW estimates of the unknown parameters and RMS. The results are reported in Table 6. As we can see from Table 6, the PINMAP(1) has the lowest RMS compared to the other models, which indicates that the PINMAP(1) provides the best forecasting for the rubella series. Figure 4 shows the plots of the rubella data series and their predicted values based on the PINMAP(1) model.

6.3. Number of Earthquakes Magnitude 8.0 to 9.9

The third example assumes the number of earthquakes magnitude 8.0 to 9.9, annually from 1977 to 2006 in the world (http://www.johnstonsarchive.net/other/quake1.html). The sample path, autocorrelation and partial autocorrelation functions are shown in Figure 5. According to Figure 5, we observe that an INMA(1) process can be suitable for modeling this data set since there exists a cut-off

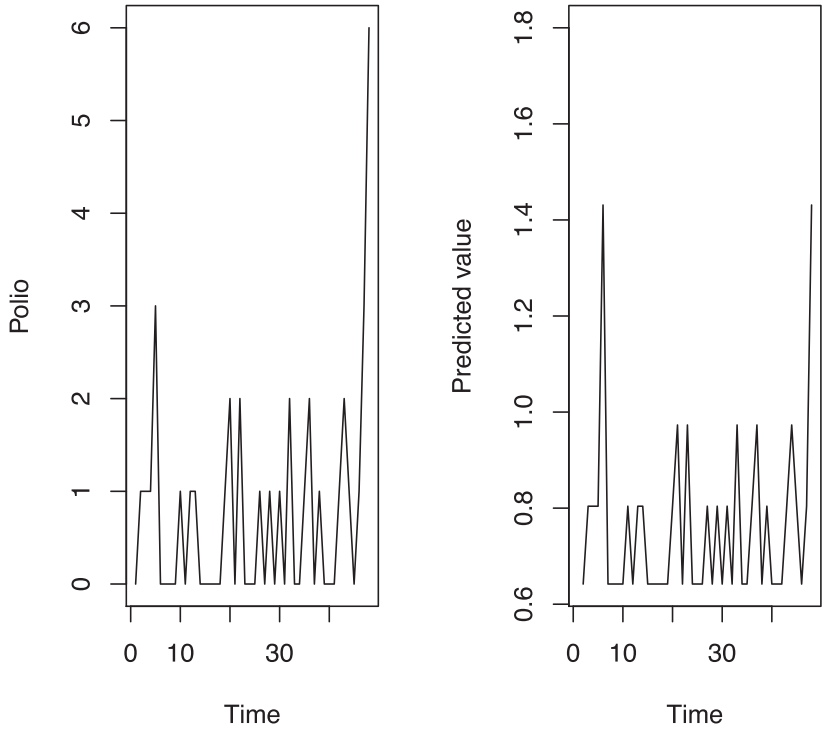

Figure 2 | Polio data and their predicted values.

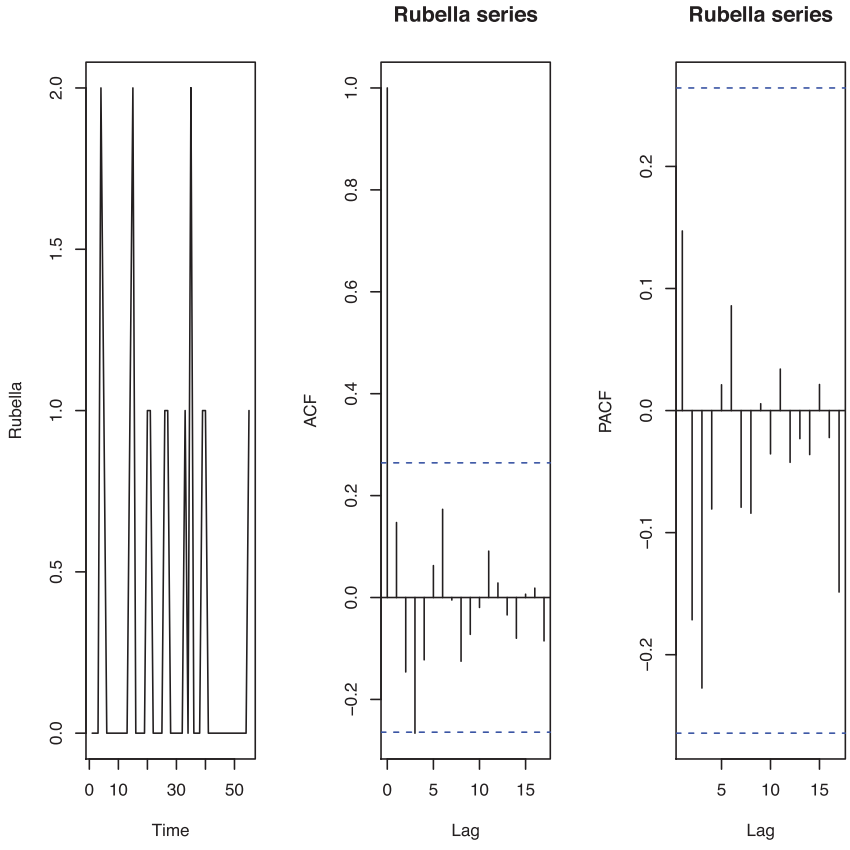

Figure 3 | The sample path, ACF and PACF plots of the number of rubella cases, monthly from Jan 2013 to Jul 2017 in Spain.

Table 6 YW estimates of the parameters and RMS for the rubella series.

Model	YW Estimates			RMS
	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{p}$	
PINMAP(1)	0.178	0.265	–	0.5586
PINMAG(1)	0.175	–	0.802	0.5588
PINMA(1)	0.172	0.274	–	0.5595
GINMA(1)	0.170	–	0.801	0.5594
NBINMAP(1)	0.368	0.128	–	0.5663
NBINMAG(1)	0.182	–	0.802	0.6111

RMS, root mean square; PINMAP(1), first-order nonnegative integer-valued moving average process with power series innovations based on a Poisson thinning operator; YW, Yule-Walker

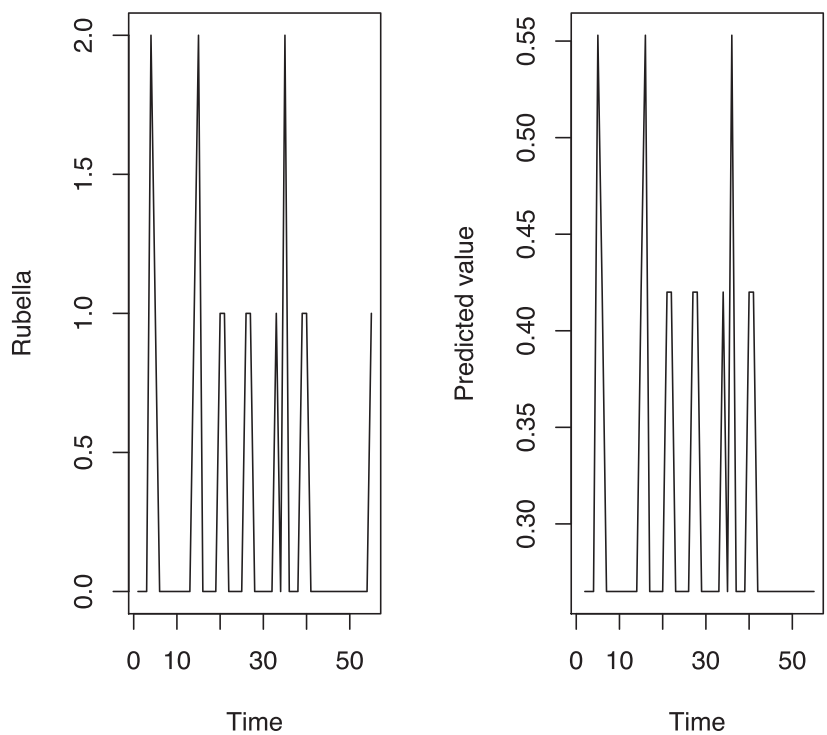

Figure 4 Rubella data and their predicted values.

Table 7 YW estimates of the parameters and RMS for the earthquakes series.

Model	YW Estimates		RMS
	$\hat{\alpha}$	$\hat{p}$	
PINMABE(1)	0.234	0.54	0.704
NBINMABE(1)	0.271	0.52	0.719

RMS, root mean square; YW, Yule-Walker

after lag 1 in the sample autocorrelation. The sample mean, variance and empirical index of dispersion are, respectively, 0.67, 0.57 and 0.86. Since the index of dispersion lower than 1, the earthquakes series is underdispersed. Thus, an underdispersed model must be assumed for modeling the series. So we fit the PINMABE(1) and NBINMABE(1) models to this data set. For both models, we obtain the YW estimates and RMS. The results are shown in Table 7. According to the values of Table 7, we conclude that the PINMABE(1) model gives better forecasting than the NBINMABE(1) model for the earthquakes series because its RMS is smaller. Figure 6 shows the plots of the earthquakes data series and their predicted values based on the PINMABE(1) model.

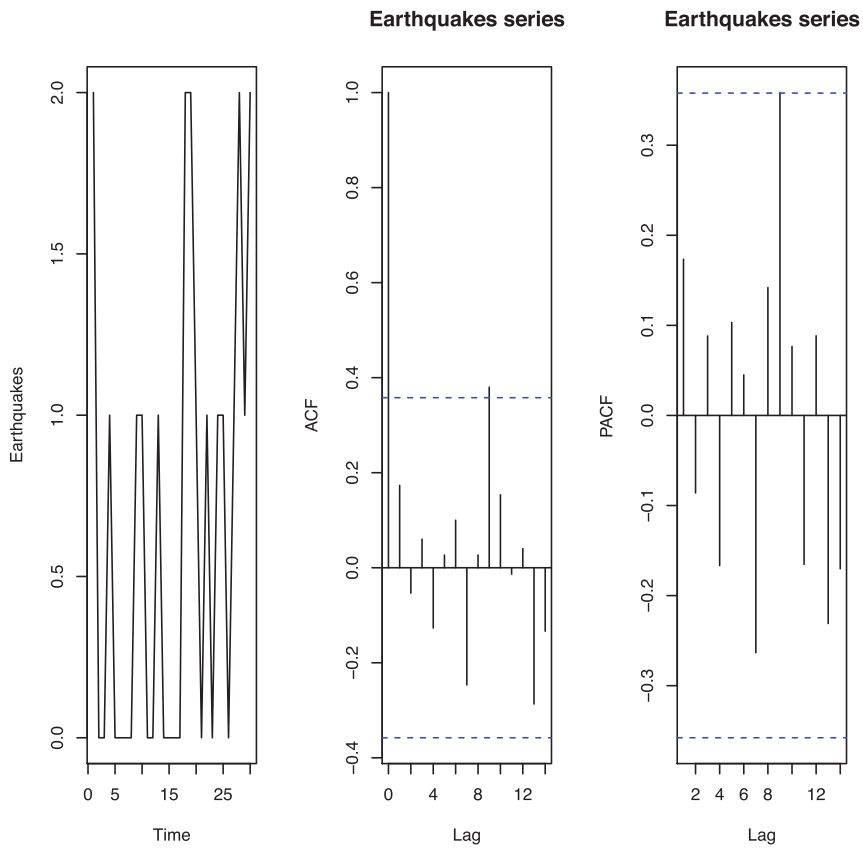

Figure 5 The sample path, ACF and PACF plots of the number of earthquakes magnitude 8.0 to 9.9, annually from 1977 to 2006 in the world.

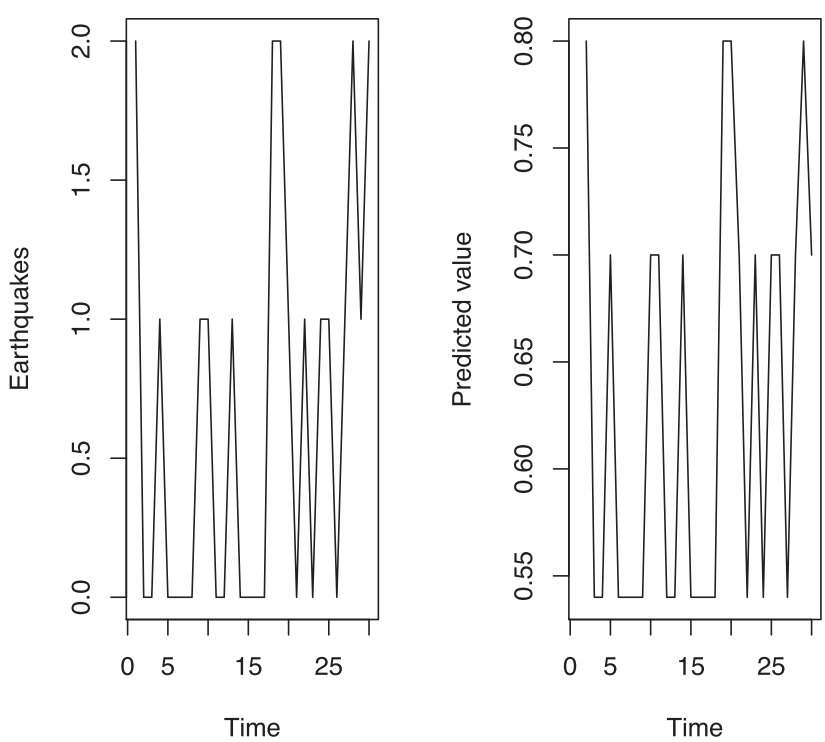

Figure 6 Earthquakes data and their predicted values.

7. CONCLUSION

In this paper, we consider a PINMAPS(1). Some statistical properties of the process are obtained. The stationary and ergodicity of the process are investigated. The parameters of the model are estimated using three methods contain YW, CLS and FGLS; also their performance is evaluated via simulation. Some sub-models are studied in detail. Finally, the model is applied to three real data sets and is shown the better performance of the model for predicting future values of overdispersed and underdispersed count data.

CONFLICTS OF INTEREST

There is no conflict of interest in this article.

AUTHORS' CONTRIBUTIONS

Mrs Rostami is the phd student under supervision of Prof. Eisa Mahmoudi and this work is related to her thesis.

ACKNOWLEDGMENTS

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 grateful to the Editor-in-Chief and Editor for their helpful remarks on improving this manuscript. The authors are also indebted to Yazd University for supporting this research.

REFERENCES

1. M.A. Al-Osh, A.A. Alzaid, J. Time Ser. Anal. 8 (1987), 261–275.
2. E. McKenzie, Adv. Appl. Probab. 20 (1988), 822–835.
3. M.A. Al-Osh, A.A. Alzaid, Stat. Pap. 29 (1988), 281–300.
4. K. Brännäs, A. Hall, Appl. Stoch. Model. Bus. Ind. 17 (2001), 277–291.
5. A.M.M. Quoreshi, Commun. Stat. Theory Methods. 35 (2006), 1343–1358.
6. M.M. Ristíc, H.S. Bakouch, A.S. Nastíc, J. Stat. Plan. Inference. 139 (2009), 2218–2226.
7. K. Brännäs, A.M.M. Quoreshi, J. Financ. Econ. 20 (2010), 1429–1440.
8. A.S. Nastíc, M.M. Ristíc, H.S. Bakouch, Math. Comput. Model. 55 (2012), 1665–1672.
9. M.M. Ristíc, A.S. Nastíc, K. Jayakumar, H.S. Bakouch, Appl. Math. Lett. 25 (2012), 481–485.
10. S. Schweer, C.H. Wei, Comput. Stat. Data Anal. 77 (2014), 267–284.
11. K. Yu, H. Zou, D. Shi, Math. Probl. Eng. 2014 (2014), 1–5.
12. M. Bourguinon, K.L.P. Vasconcellos, Braz. J. Probab. Stat. 29 (2015), 71–93.
13. K. Yu, H. Zou, J. Appl. Math. 2015 (2015), 1–7.
14. M. Mohammadpour, H.S. Bakouch, M. Shirozhan, Braz. J. Probab. Stat. 32 (2018), 262–280.
15. F.W. Steutel, K. van Harn, Ann. Probab. 7 (1979), 893–899.
16. A.A. Alzaid, M.A. Al-Osh, Stat. Neerl. 42 (1988), 53–61.
17. S. Kocherlakota, K. Kocherlakota, Bivariate Discrete Distributions, Statistics: Textbooks and Monographs, vol. 132, Markel Dekker, New York, 1992.