Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2015, Article ID 457842, 7 pages
http://dx.doi.org/10.1155/2015/457842

Research Article

The Combined Poisson INMA(q) Models for Time Series of Counts

Kaizhi Yu[1] and Hong Zou[2]

[1] *Statistics School, Southwestern University of Finance and Economics, Chengdu 611130, China*
[2] *School of Economics, Southwestern University of Finance and Economics, Chengdu 611130, China*

Correspondence should be addressed to Kaizhi Yu; kaizhiyu.swufe@gmail.com

Received 11 October 2014; Revised 14 January 2015; Accepted 14 January 2015

Academic Editor: Angelo Ciaramella

Copyright © 2015 K. Yu and H. Zou. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A new stationary qth-order integer-valued moving average process with Poisson innovation is introduced based on decision random vector. Some statistical properties of the process are established. Estimators of the parameters of the process are obtained using the method of moments. Some numerical results of the estimators are presented to assess the performance of moment estimators.

1. Introduction

In natural and social sciences, time series of correlated counting are met very often. In particular, in economics and medicine many interesting variables are nonnegative count data, for example, the number of shareholders in large Finnish and Swedish stocks, big numbers even for frequently traded stocks, the number of arrivals per week to the emergency service of the hospital, and monthly polio incidence counts in Germany. Most of the research on count processes assumes that the count data are independent and identically distributed. However, in practice, observations may be autocorrelated, and this may adversely affect the performance of traditional model developed under the assumption of independence. In recent years, count data time series models have been devised to avoid making restrictive assumptions on the distribution of the error term. Regression models for time series count data have been proposed [1, 2]. On the other hand, there have been attempts to develop suitable classes of models that resemble the structure and properties of the usual linear ARMA models. For instances, Al-Osh and Alzaid [3] proposed an integer-valued moving average model (INMA) for discrete data. An integer-valued GARCH model was given to study overdispersed counts [4]. Silva et al. [5] considered the problem of forecasting in INAR(1) model. Random coefficient INAR models were introduced by Zheng et al. [6, 7]. The signed thinning operator was developed by Kachour and Truquet [8]. A new stationary first-order integer-valued autoregressive process with geometric marginal distribution based on the generalized binomial thinning was introduced by Ristić et al. [9]. In this analysis of counts, the class of integer-valued moving average models plays an important role.

The nonnegative integer-valued moving average process of the order q (INMAR(q)) was introduced by Al-Osh and Alzaid [3]. The INMA(q) process is defined by the recursion,

$$X_t = \theta_1 \circ_t \varepsilon_{t-1} + \cdots + \theta_q \circ_t \varepsilon_{t-q} + \varepsilon_t, \tag{1}$$

where $\theta_k \circ_t \varepsilon_{t-k} = \sum_{i=1}^{\varepsilon_{t-k}} \nu_i$ and the $\{\nu_i\}$ designated by counting series is a sequence of i.i.d. Bernoulli random variables with $E(\nu_i) = \theta$, independent of ε_{t-k}. The parameters $\theta_1, \dots, \theta_q \in (0, 1)$. The thinning operation "$\circ_t$" indicates the corresponding thinning is associated with time. The terms $\theta_i \circ_t \varepsilon_t$ and $\theta_j \circ_t \varepsilon_t$ are independent. The choice of appropriate marginal distributions is still problematic for getting a particular distribution of X_t. To overcome these difficulties, Weiß [10] introduced combined INAR(p) models by using "decision" random variables. Ristić et al. [11] considered piecewise functions for count data model. Therefore, we adopt the similar approach to deal with the problem in INMA(q) models. In this paper, we propose a combined INMA(q) model by allowing the parameters value to vary with "decision" random vector.

The paper is organized as follows. In Section 2, we specify the model and derive some statistical properties. Section 3 concerns unknown parameter estimation by Yule-Walker method. In Section 4, we conduct some Monte Carlo simulations. Finally, Section 5 concludes.

2. Definition and Basic Properties of the PCINMA(q) Process

Definition 1 (PCINMA(q) model). A stochastic process $\{X_t\}$ is said to be the Poisson combined INMA(q) process if it satisfies the following recursive equations:

$$X_t = D_{t,1}\left(\theta\circ_t\varepsilon_{t-1}\right) + \cdots + D_{t,q}\left(\theta\circ_t\varepsilon_{t-q}\right) + \varepsilon_t, \qquad (2)$$

where $\{\varepsilon_t; t \in \mathbb{Z}\}$ is a sequence of independent and identically distributed Poisson random variables with parameter λ and $\theta \in (0,1)$. $\{\mathbf{D}_t; t \in \mathbb{Z}\}$ is an i.i.d. process of "decision" random vector $\mathbf{D}_t = (D_{t,1},\dots,D_{t,q}) \sim \mathrm{Mult}(1;\phi_1,\dots,\phi_q)$, independent of $\{\varepsilon_t; t \in \mathbb{Z}\}$. Moreover, the counting series $\theta\circ_t\varepsilon_{t-k}$, $k = 1,\dots,q$, are independent of $\{\mathbf{D}_t\}$ and ε_t at time t. To our knowledge, few efforts have been devoted to studying combined INMA models. In this paper, we aim to fill this gap. Definition 1 shows that X_t is equal to $\theta\circ_t\varepsilon_{t-1} + \varepsilon_t$ with probability ϕ_1 or equal to $\theta\circ_t\varepsilon_{t-2} + \varepsilon_t$ with probability $\phi_2,\dots,$ or else equal to $\theta\circ_t\varepsilon_{t-q} + \varepsilon_t$ with probability ϕ_q $(= 1-\phi_1-\phi_2-\cdots-\phi_{q-1})$.

The moments will be useful in obtaining the appropriate estimating equations for parameter estimation.

Theorem 2. *The numerical characteristics of $\{X_t\}$ are as follows:*

(i)

$$\mu_X := E\left(X_t\right) = \lambda\left(\theta+1\right), \qquad (3)$$

(ii)

$$\sigma_X^2 := \mathrm{Var}\left(X_t\right) = \theta\left(1-\theta\right)\lambda + \left(\theta^2+1\right)\lambda^2, \qquad (4)$$

(iii)

$$\gamma_X(k) := \mathrm{cov}\left(X_t, X_{t-k}\right) = \begin{cases} \theta\lambda^2\phi_k + \theta^2\lambda^2 \sum_{j=k+1}^{q}\phi_j\phi_{j-k}, & 1 \le k \le q-1, \\ \theta\lambda^2\phi_k, & k = q, \\ 0, & k \ge q+1. \end{cases} \qquad (5)$$

Proof. (i) Consider

$$\begin{aligned} E\left(X_t\right) &= E\left[D_{t,1}\left(\theta\circ_t\varepsilon_{t-1}+\varepsilon_t\right)\right] + \cdots + E\left[D_{t,q}\left(\theta\circ_t\varepsilon_{t-q}+\varepsilon_t\right)\right] \\ &= E\left(D_{t,1}\right)E\left(\theta\circ_t\varepsilon_{t-1}+\varepsilon_t\right) + \cdots + E\left(D_{t,q}\right) \\ &\quad \cdot E\left(\theta\circ_t\varepsilon_{t-q}+\varepsilon_t\right) \\ &= \phi_1\left(\lambda\theta+\lambda\right) + \cdots + \phi_q\left(\lambda\theta+\lambda\right) = \lambda\left(\theta+1\right). \end{aligned} \qquad (6)$$

(ii) Moreover,

$$\begin{aligned} \mathrm{Var}\left(X_t\right) &= \mathrm{Var}\left(\sum_{j=1}^{q} D_{t,j}\left(\theta\circ_t\varepsilon_{t-j}+\varepsilon_t\right)\right) \\ &= \sum_{j=1}^{q}\mathrm{Var}\left(D_{t,j}\left(\theta\circ_t\varepsilon_{t-j}+\varepsilon_t\right)\right) \\ &\quad + 2\sum_{i<j}\mathrm{cov}\left(D_{t,i}\left(\theta\circ_t\varepsilon_{t-i}+\varepsilon_t\right), D_{t,j}\left(\theta\circ_t\varepsilon_{t-j}+\varepsilon_t\right)\right). \end{aligned} \qquad (7)$$

Note that $\mathrm{Var}(D_{t,j}(\theta\circ_t\varepsilon_{t-j}+\varepsilon_t))$ in first summation on the right-hand side is

$$\begin{aligned} &= \mathrm{Var}\left(D_{t,j}\right)E\left(\theta\circ_t\varepsilon_{t-j}+\varepsilon_t\right)^2 + E^2\left(D_{t,j}\right)\mathrm{Var}\left(\theta\circ_t\varepsilon_{t-j}+\varepsilon_t\right) \\ &= \phi_j\left(1-\phi_j\right)E\left(\left(\theta\circ_t\varepsilon_{t-j}\right)^2 + \varepsilon_t^2 + 2\left(\theta\circ_t\varepsilon_{t-j}\right)\varepsilon_t\right) \\ &\quad + \phi_j^2\left(\mathrm{Var}\left(\theta\circ_t\varepsilon_{t-j}\right) + \mathrm{Var}\left(\varepsilon_t\right) + 2\,\mathrm{cov}\left(\theta\circ_t\varepsilon_{t-j},\varepsilon_t\right)\right) \\ &= \phi_j\left(1-\phi_j\right)\left\{\left[\theta^2E\left(\varepsilon_{t-j}^2\right) + \theta\left(1-\theta\right)E\left(\varepsilon_{t-j}\right)\right] + E\left(\varepsilon_t^2\right)\right. \\ &\qquad \left. + 2E\left(\theta\circ_t\varepsilon_{t-j}\right)E\left(\varepsilon_t\right)\right\} \\ &\quad + \phi_j^2\left\{\left[\theta^2\,\mathrm{Var}\left(\varepsilon_{t-j}\right) + \theta\left(1-\theta\right)E\left(\varepsilon_{t-j}\right)\right] + \sigma_\varepsilon^2\right\} \\ &= \phi_j\left(1-\phi_j\right)\left[\left(\theta+1\right)^2\lambda^2 + \theta\left(1-\theta\right)\lambda + \left(\theta^2+1\right)\lambda^2\right] \\ &\quad + \phi_j^2\left[\left(\theta^2+1\right)\lambda^2 + \theta\left(1-\theta\right)\lambda\right] \\ &= \phi_j\left(1-\phi_j\right)\left(\theta+1\right)^2\lambda^2 + \phi_j\left[\theta\left(1-\theta\right)\lambda + \left(\theta^2+1\right)\lambda^2\right]. \end{aligned} \qquad (8)$$

By the same arguments as above, it follows that

$$\begin{aligned} &\mathrm{cov}\left(D_{t,i}\left(\theta\circ_t\varepsilon_{t-i}+\varepsilon_t\right), D_{t,j}\left(\theta\circ_t\varepsilon_{t-j}+\varepsilon_t\right)\right) \\ &\quad = \mathrm{cov}\left(D_{t,i}\left(\theta\circ_t\varepsilon_{t-i}\right), D_{t,j}\left(\theta\circ_t\varepsilon_{t-j}\right)\right) \\ &\qquad + \mathrm{cov}\left(D_{t,i}\left(\theta\circ_t\varepsilon_{t-i}\right), D_{t,j}\varepsilon_t\right) \\ &\qquad + \mathrm{cov}\left(D_{t,i}\varepsilon_t, D_{t,j}\left(\alpha\circ_t\varepsilon_{t-j}\right)\right) + \mathrm{cov}\left(D_{t,i}\varepsilon_t, D_{t,j}\varepsilon_t\right) \\ &\quad = \mathrm{cov}\left(D_{t,i}, D_{t,j}\right)E\left(\theta\circ_t\varepsilon_{t-i}\right)E\left(\theta\circ_t\varepsilon_{t-j}\right) \\ &\qquad + \mathrm{cov}\left(D_{t,i}, D_{t,j}\right)E\left(\theta\circ_t\varepsilon_{t-i}\right)E\left(\varepsilon_t\right) \\ &\qquad + \mathrm{cov}\left(D_{t,i}, D_{t,j}\right)E\left(\varepsilon_t\right)E\left(\theta\circ_t\varepsilon_{t-j}\right) \\ &\qquad + \mathrm{cov}\left(D_{t,i}, D_{t,j}\right)E^2\left(\varepsilon_t\right) \\ &\quad = -\phi_i\phi_j\left(\theta^2\lambda^2 + \theta\lambda^2 + \theta\lambda^2 + \lambda^2\right) = -\phi_i\phi_j\left(\theta+1\right)^2\lambda^2. \end{aligned} \qquad (9)$$

Therefore,

$$\begin{aligned}
\operatorname{Var}(X_t) &= \operatorname{Var}\left(\sum_{j=1}^{q} D_{t,j}\left(\theta\circ_t\varepsilon_{t-j}+\varepsilon_t\right)\right)\\
&= \sum_{j=1}^{q}\left(\phi_j\left(1-\phi_j\right)(\theta+1)^2\lambda^2\right.\\
&\qquad\left.+\phi_j\left[\theta(1-\theta)\lambda+\left(\theta^2+1\right)\lambda^2\right]\right)\\
&\quad-2\sum_{i<j}\phi_i\phi_j(\theta+1)^2\lambda^2\\
&= (\theta+1)^2\lambda^2\left(1-\sum_{j=1}^{q}\phi_j^2\right)\\
&\quad+\left[\theta(1-\theta)\lambda+\left(\theta^2+1\right)\lambda^2\right]\\
&\quad-2(\theta+1)^2\lambda^2\sum_{i<j}\phi_i\phi_j\\
&= \theta(1-\theta)\lambda+\left(\theta^2+1\right)\lambda^2.
\end{aligned}\tag{10}$$

(iii) For $1 \le k \le q-1$, the autocovariance function of X_t is

$$\begin{aligned}
&\operatorname{cov}(X_t, X_{t-k})\\
&= \operatorname{cov}\left(\sum_{i=1}^{q} D_{t,i}(\theta\circ_t\varepsilon_{t-i}+\varepsilon_t),\right.\\
&\qquad\left.\sum_{j=1}^{q} D_{t-k,j}\left(\theta\circ_{t-k}\varepsilon_{t-k-j}+\varepsilon_{t-k}\right)\right)\\
&= \operatorname{cov}\left(D_{t,k}(\theta\circ_t\varepsilon_{t-k}+\varepsilon_t),\right.\\
&\qquad\left.\sum_{j=1}^{q} D_{t-k,j}\left(\theta\circ_{t-k}\varepsilon_{t-k-j}+\varepsilon_{t-k}\right)\right)\\
&\quad+\sum_{j=k+1}^{q}\operatorname{cov}\left(D_{t,j}\left(\theta\circ_{t-j}\varepsilon_{t-j}+\varepsilon_t\right),\right.\\
&\qquad\left.D_{t-k,j-k}\left(\theta\circ_{t-k}\varepsilon_{t-k-(j-k)}+\varepsilon_{t-k}\right)\right)\\
&= \sum_{j=1}^{q}\operatorname{cov}\left(D_{t,k}(\theta\circ_t\varepsilon_{t-k}), D_{t-k,j}\varepsilon_{t-k}\right)\\
&\quad+\sum_{j=k+1}^{q}\operatorname{cov}\left(D_{t,j}\left(\theta\circ_{t-j}\varepsilon_{t-j}\right), D_{t-k,j-k}\left(\theta\circ_{t-k}\varepsilon_{t-j}\right)\right)\\
&= \sum_{j=1}^{q}\operatorname{cov}(\theta\circ_t\varepsilon_{t-k}, \varepsilon_{t-k})E(D_{t,k})E\left(D_{t-k,j}\right)\\
&\quad+\sum_{j=k+1}^{q}\operatorname{cov}\left(\theta\circ_{t-j}\varepsilon_{t-j}, \theta\circ_{t-k}\varepsilon_{t-j}\right)E\left(D_{t,j}\right)E\left(D_{t-k,j-k}\right)\\
&= \sum_{j=1}^{q}\theta\operatorname{Var}(\varepsilon_{t-k})\phi_k\phi_j+\sum_{j=k+1}^{q}\theta^2\operatorname{Var}\left(\varepsilon_{t-j}\right)\phi_j\phi_{j-k}\\
&= \theta\lambda^2\phi_k+\theta^2\lambda^2\sum_{j=k+1}^{q}\phi_j\phi_{j-k}.
\end{aligned}\tag{11}$$

If $k = q$, by using a similar approach, we get $\operatorname{cov}(X_t, X_{t-q}) = \theta\lambda^2\phi_q$. For $k \ge q+1$, all the terms, $\theta\circ_t\varepsilon_{t-1},\dots,\theta\circ_t\varepsilon_{t-q},\varepsilon_t,\theta\circ_{t-k}\varepsilon_{t-k-1},\dots,\theta\circ_{t-k}\varepsilon_{t-k-q}$ and ε_{t-k} involved in X_t and X_{t-k}, are mutually independent. Therefore, the autocovariance function of X_t is equal to zero for $k \ge q+1$. □

Theorem 3. *Let X_t be the process defined by the equation in (2). Then*

(i) *$\{X_t\}$ is a covariance stationary process;*

(ii) *$E(X_t^k) \le C < \infty$, $k = 1,2,3,4$, for some constant $C > 0$.*

Proof. (i) The first conclusion is immediate from the definition of covariance stationary process.

(ii) For $k = 1$, it is straightforward.

For $k = 2$, it follows that

$$\begin{aligned}
E\left(X_t^2\right) \le \max\Big\{&E\left(D_{t,1}(\theta\circ_t\varepsilon_{t-1}+\varepsilon_t)\right)^2,\dots,\\
&E\left(D_{t,q}\left(\theta\circ_t\varepsilon_{t-q}+\varepsilon_t\right)\right)^2\Big\}.
\end{aligned}\tag{12}$$

Note that, for $j = 1,2,\dots,q$,

$$\begin{aligned}
&E\left(D_{t,j}\left(\theta\circ_t\varepsilon_{t-j}+\varepsilon_t\right)\right)^2\\
&= E\left(D_{t,j}\right)^2E\left(\theta\circ_t\varepsilon_{t-j}+\varepsilon_t\right)^2\\
&= E\left(D_{t,j}\right)^2\left(E\left(\theta\circ_t\varepsilon_{t-j}\right)^2+E(\varepsilon_t)^2+2E\left[\left(\theta\circ_t\varepsilon_{t-j}\right)\varepsilon_t\right]\right)\\
&= \phi_j\left(1-\phi_j\right)\\
&\quad\cdot\left[\theta^2\left(\lambda+\lambda^2\right)+\theta(1-\theta)\lambda+\left(\lambda+\lambda^2\right)+2\theta\lambda^2\right]\\
&= \phi_j\left(1-\phi_j\right)(\theta+1)\lambda\left[(\theta+1)\lambda+1\right].
\end{aligned}\tag{13}$$

Then we have

$$E\left(X_t^2\right) \le \frac{1}{4}(\theta+1)\lambda\left[(\theta+1)\lambda+1\right] < \infty. \tag{14}$$

Similarly, for $k = 3$,

$$\begin{aligned}
E\left(X_t^3\right) \le \max\Big\{&E\left(D_{t,1}(\theta\circ_t\varepsilon_{t-1}+\varepsilon_t)\right)^3,\dots,\\
&E\left(D_{t,q}\left(\theta\circ_t\varepsilon_{t-q}+\varepsilon_t\right)\right)^3\Big\}.
\end{aligned}\tag{15}$$

Note that, for $j = 1, 2, \ldots, q$,

$$\begin{aligned}
&E\left(D_{t,j}\left(\theta\circ_t\varepsilon_{t-j}+\varepsilon_t\right)\right)^3 \\
&\quad = E\left(D_{t,j}\right)^3 E\left(\theta\circ_t\varepsilon_{t-j}+\varepsilon_t\right)^3 \\
&\quad = E\left(D_{t,j}\right)^3 \left(E\left(\theta\circ_t\varepsilon_{t-j}\right)^3 + E\left(\varepsilon_t\right)^3\right. \\
&\qquad\qquad \left. + 3E\left[\left(\theta\circ_t\varepsilon_{t-j}\right)^2\varepsilon_t\right] + E\left[\left(\theta\circ_t\varepsilon_{t-j}\right)\varepsilon_t^2\right]\right) \\
&\quad = \phi_j^3\left[\theta^3\tau + 3\theta^2(1-\theta)\left(\lambda+\lambda^2\right)\right. \\
&\qquad\qquad \left. + \left(\theta - 3\theta^2(1-\theta) - \theta^3\right)\lambda\right] \\
&\qquad + \phi_j^3\tau + 3\phi_j^3\left\{\left[\theta^2\left(\lambda+\lambda^2\right) + \theta(1-\theta)\lambda\right]\lambda\right\} \\
&\qquad + 3\phi_j^3\theta\lambda\left(\lambda+\lambda^2\right) \\
&\quad = \phi_j^3\left\{\theta\lambda\left[\theta^2(\tau-1-3\lambda) + 3\lambda(1+\lambda)\theta + 3\lambda^2 + 6\lambda + 1\right]\right. \\
&\qquad\qquad \left. + \tau\right\},
\end{aligned} \tag{16}$$

where $\tau := \lambda^3 + 3\lambda^2 + \lambda$. Let $\phi^3_{\max} = \max(\phi_1^3, \ldots, \phi_q^3)$. Then

$$\begin{aligned}
&E\left(X_t^3\right) \\
&\quad \le \phi^3_{\max}\left\{\theta\lambda\left[\theta^2(\tau-1-3\lambda) + 3\lambda(1+\lambda)\theta + 3\lambda^2\right.\right. \\
&\qquad\qquad \left.\left. + 6\lambda + 1\right] + \tau\right\} < \infty.
\end{aligned} \tag{17}$$

After some tedious calculations, we also can show the result holds for $k = 4$. We skip the details. Next, we will present ergodic theorem for stationary process $\{X_t\}$. There are a variety of ergodic theorems, differing in their assumptions and in the modes of convergence. Here the convergence is in mean square. The next two lemmas will be useful in proving the ergodicity of the sample mean and sample autocovariance function of $\{X_t\}$. □

Lemma 4. *If $\{Z_t\}$ is stationary with mean μ_Z and autocovariance function $\gamma_Z(\cdot)$, then as $T \to \infty$, $\operatorname{Var}(\overline{Z}_T) = E(\overline{Z}_T - \mu_Z)^2 \to 0$, if $\gamma_Z(T) \to 0$, where $\overline{Z}_T := (1/T)\sum_{t=1}^{T} X_t$.*

Proof. See Theorem 7.1.1 in Brockwell and Davis [13]. □

Lemma 5. *Suppose $\{Z_t\}$ is a covariance stationary process having covariance function $\gamma_Z(v) := E(Z_{t+v}Z_t)$ and a mean of zero. If $\lim_{T\to\infty}(1/T)\sum_{l=1}^{T}[E(Z_tZ_{t+v}Z_{t+l}Z_{t+l+v}) - \gamma_Z^2(v)] = 0$. Then, for any fixed $v = 0, 1, 2, \ldots$,*

$$\lim_{T\to\infty}\frac{1}{T}\sum_{l=1}^{T} E\left[\widehat{\gamma}_Z(v) - \gamma_Z^2(v)\right]^2 = 0, \tag{18}$$

where $\widehat{\gamma}_Z(v)$ is the sample covariance function $\widehat{\gamma}_Z(v) := (1/T)\sum_{t=1}^{T-v} Z_{t+v}Z_t$.

Proof. See Theorem 5.2 in Karlin and Taylor [12]. □

Lemma 6. *Let process $Y_t := X_t - \mu_X$ be a transformation of X_t; then the following results hold:*

(i) *$\{Y_t\}$ is a covariance stationary process with zero mean;*
(ii) *$\gamma_Y(k) := \operatorname{cov}(Y_t, Y_{t-k}) = \gamma_X(k)$, $k = 1, 2, 3, \ldots$;*
(iii) *$E(Y_t^k) \le C^* < \infty$, $k = 1, 2, 3, 4$, for some constant $C^* > 0$;*
(iv) *$\{Y_t\}$ is ergodic in autocovariance function.*

Proof. The only part of this lemma that is not obvious is part (iv). The proof of properties (iv) is as follows.

To prove that $\{Y_t\}$ is ergodic in autocovariance function, it suffices to show that

$$\lim_{T\to\infty}\frac{1}{T}\sum_{l=1}^{T}\left[E\left(Y_tY_{t+v}Y_{t+l}Y_{t+l+v}\right) - \gamma_Y^2(v)\right] = 0 \tag{19}$$

according to Lemma 5. Thus, we will discuss two cases. For simplicity in notation, we define $c_2 := E(Y_t^2)$, $c_3 := E(Y_t^3)$, $c_4 := E(Y_t^4)$, and $R_Y(v) := E(Y_tY_{t+v})$.

Case 1. For $v = 0$, $E(Y_tY_{t+v}Y_{t+l}Y_{t+l+v}) = E(Y_t^2Y_{t+l}^2)$, $\gamma_Y^2(v) = \gamma_X^2(v) = \sigma_X^2$.

If $1 \le l \le q$, using the Schwarz inequality, we get

$$E\left(Y_t^2Y_{t+l}^2\right) \le \sqrt{E\left(\left|Y_t^2\right|^2\right)E\left(\left|Y_{t+l}^2\right|^2\right)} = c_4 \le C^*. \tag{20}$$

If $l \ge q+1$, note that Y_{t+l} and Y_t are irrelevant, and then

$$E\left(Y_t^2Y_{t+l}^2\right) = E\left(Y_t^2\right)E\left(Y_{t+l}^2\right) = c_2^2 = \gamma_X^2(v). \tag{21}$$

Therefore, $(1/T)\sum_{l=1}^{T}[E(Y_tY_{t+v}Y_{t+l}Y_{t+l+v}) - \gamma_Y^2(v)] \le (1/T)[2(c_4 - \sigma_X^2)] \to 0$, for $T \to \infty$.

Case 2. For $v \ge 1$, $R_Y^2(v) = [E(Y_tY_{t+v})]^2 = \gamma_X^2(v)$.

If $1 \le l \le q+v$, using Schwarz inequality twice,

$$\begin{aligned}
&E\left(Y_tY_{t+v}Y_{t+l}Y_{t+l+v}\right) \\
&\quad \le \sqrt{E\left(\left|Y_tY_{t+v}\right|^2\right)E\left(\left|Y_{t+l}Y_{t+l+v}\right|^2\right)} \\
&\quad \le \sqrt[1/4]{E\left(Y_t^4\right)E\left(Y_{t+v}^4\right)E\left(Y_{t+l}^4\right)E\left(Y_{t+l+v}^4\right)} = c_4.
\end{aligned} \tag{22}$$

If $l \ge q+v+1$, note that X_tX_{t+v}与$X_{t+l}X_{t+l+v}$ uncorrelation; thus, we have

$$E\left(Y_tY_{t+v}Y_{t+l}Y_{t+l+v}\right) = E\left(X_tX_{t+v}\right)E\left(Y_{t+l}Y_{t+l+v}\right) = R_Y^2(v). \tag{23}$$

Then $(1/T)\sum_{l=1}^{T}[E(Y_tY_{t+v}Y_{t+l}Y_{t+l+v}) - \gamma_Y^2(v)] \le (1/T)\{(2+v)[c_4 - R_Y^2(v)]\} \to 0$, for $T \to \infty$.

This proves Lemma 6. □

Theorem 7. *Let $\{X_t\}$ be a PCINMA(q) process according to Definition 1. Then, the stochastic process $\{X_t\}$ is ergodic in the mean and autocovariance function.*

Proof. For notational simplicity, we define $\overline{X}_T(k) := (1/T)\sum_{t=1}^{T-k} X_{t+k}$, $\widehat{\gamma}_Y(k) := (1/T)\sum_{t=1}^{T-k} Y_{t+k}Y_t$, and $\widehat{\gamma}_X(k) := (1/T)\sum_{t=1}^{T-k}(X_{t+k} - \overline{X}_T)(X_t - \overline{X}_T)$. And we assume here that the sample consists of $T + k$ observations on X.

(i) Note that $\gamma_X(k) \to 0$, for $k \to \infty$. From the result of Lemma 4, we obtain $\mathrm{Var}(\overline{X}_T) = E(\overline{X}_T - \mu_X)^2 \to 0$.

Then $\overline{X}_T$ converges in probability to μ_X. Therefore, the process $\{X_t\}$ is ergodic in the mean.

Next, we prove that the $\{X_t\}$ is ergodic for second moment by induction.

(ii) First we prove $\overline{X}_T(k) \xrightarrow{P} \mu_X$. Suppose $\varepsilon_1 > 0$ is given:

$$\begin{aligned} P\left(\left|\overline{X}_T(k) - \mu_X\right| \geq \varepsilon_1\right) &\leq P\left(\left|\overline{X}_T(k) - \overline{X}\right| \geq \frac{\varepsilon_1}{2}\right) + P\left(\left|\overline{X}_T - \mu_X\right| \geq \frac{\varepsilon_1}{2}\right) \\ &= P\left(\left|\frac{1}{T}\sum_{t=1}^{k} X_t\right| \geq \frac{\varepsilon_1}{2}\right) + P\left(\left|\overline{X}_T - \mu_X\right| \geq \frac{\varepsilon_1}{2}\right) \\ &\leq \sum_{t=1}^{k} P\left(\frac{1}{T}\left|X_t\right| \geq \frac{\varepsilon_1}{2}\right) + P\left(\left|\overline{X}_T - \mu_X\right| \geq \frac{\varepsilon_1}{2}\right). \end{aligned} \tag{24}$$

Using the Markov inequality, $\sum_{t=1}^{k} P((1/T)|X_t| \geq \varepsilon_1/2) \leq \sum_{t=1}^{k}(E(X_t)/(1/2)T\varepsilon_1) \to 0$, for $T \to \infty$.

Since $\{X_t\}$ is ergodic in the mean, thus $P(|\overline{X}_T - \mu_X| \geq \varepsilon_1/2) \to 0$, for $T \to \infty$.

Therefore, $\overline{X}_T(k) \xrightarrow{P} \mu_X$, for $T \to \infty$.

Now we prove the second result $\widehat{\gamma}_X(k) - \gamma_X(k) \xrightarrow{P} 0$.

Consider any $\varepsilon > 0$:

$$\begin{aligned} P\left(\left|\widehat{\gamma}_X(k) - \gamma_X(k)\right| \geq \varepsilon\right) &= P\left(\left|\widehat{\gamma}_X(k) - \widehat{\gamma}_Y(k) + \widehat{\gamma}_Y(k) - \gamma_X(k)\right| \geq \varepsilon\right) \\ &\leq P\left(\left|\widehat{\gamma}_X(k) - \widehat{\gamma}_Y(k)\right| \geq \frac{\varepsilon}{2}\right) + P\left(\left|\widehat{\gamma}_Y(k) - \gamma_X(k)\right| \geq \frac{\varepsilon}{2}\right). \end{aligned} \tag{25}$$

Note that

$$\begin{aligned} P\left(\left|\widehat{\gamma}_X(k) - \widehat{\gamma}_Y(k)\right| \geq \frac{\varepsilon}{2}\right) &= P\left(\left|\left(\overline{X}_T(k) + \overline{X}_T\right)\left(\overline{X}_T - \mu_X\right) + \left(\overline{X}_T^2 - \mu_X^2\right)\right| \geq \frac{\varepsilon}{2}\right) \\ &\leq P\left(\left|\left(\overline{X}_T(k) + \overline{X}_T\right)\left(\overline{X}_T - \mu_X\right)\right| \geq \frac{\varepsilon}{4}\right) + P\left(\left|\overline{X}_T^2 - \mu_X^2\right| \geq \frac{\varepsilon}{4}\right). \end{aligned} \tag{26}$$

Since the sample mean $\overline{X}_T$ converges in probability to μ_X, according to Slutsky's theorem, we get $(\overline{X}_T(k) + \overline{X}_T)(\overline{X}_T - \mu_X) \xrightarrow{P} 0$, $\overline{X}_T^2 - \mu_X^2 \xrightarrow{P} 0$.

Then we have $P(|\widehat{\gamma}_X(k) - \widehat{\gamma}_Y(k)| \geq \varepsilon/2) \xrightarrow{P} 0$, for $T \to \infty$. From the (iv) of Lemma 6, we obtain

$$\widehat{\gamma}_Y(k) - \gamma_X(k) = \widehat{\gamma}_Y(k) - \gamma_Y(k) \xrightarrow{P} 0, \quad \text{for } T \to \infty. \tag{27}$$

And consequently, $P(|\widehat{\gamma}_X(k) - \gamma_X(k)| \geq \varepsilon) \xrightarrow{P} 0$, for $T \to \infty$.

This leads to the desired conclusion. □

3. Estimation of the Unknown Parameters

In this section, we discuss approaches to the estimation of the unknown parameters. And we assume we have T observations, $X_1, \ldots, X_T$, from a Poisson combined INMA(q) process in which the order parameter q is known. One of the main interests in the literature of INMA process is to estimate the unknown parameters. Using the sample covariance function, we get the estimators of unknown parameters $(\phi_1, \ldots, \phi_q, \theta, \lambda)$ through solving the following equations:

$$\begin{gathered} \widehat{\gamma}(0) - \left[\theta(1-\theta)\lambda + \left(\theta^2 + 1\right)\lambda^2\right] = 0 \\ \widehat{\gamma}(1) - \left(\theta\lambda^2\phi_1 + \theta^2\lambda^2\sum_{j=2}^{q}\phi_j\phi_{j-1}\right) = 0 \\ \vdots \\ \widehat{\gamma}(q-1) - \left(\theta\lambda^2\phi_{q-1} + \theta^2\lambda^2\phi_q\phi_1\right) = 0 \\ \widehat{\gamma}(q) - \theta\lambda^2\phi_q = 0. \end{gathered} \tag{28}$$

The idea behind these estimators is that of equating population moments to sample moments and then solving for the parameters in terms of the sample moments. These estimators are typically called the Yule-Walker estimators. As the $\widehat{\gamma}(k)$ consistently estimates the true autocovariance function $\gamma(k)$ [13], the Yule-Walker estimators are consistent. Following Brockwell and Davis (2009) and Billingsley [14], it is easy to show that under some mild moment conditions the marginal mean estimator $\overline{X}_T$ is asymptotically normally distributed.

4. Monte Carlo Simulation Study

We provide some simulations results to show the empirical performance of these estimators. Owing to the nonlinearity, the estimator expressions of unknown parameters are quite complicated. The aim of simulation study is to assess the finite sample performances of the moments estimators. Consider the following model:

$$X_t = D_{t,1}\left(\theta \circ_t \varepsilon_{t-1}\right) + D_{t,2}\left(\theta \circ_t \varepsilon_{t-2}\right) + D_{t,3}\left(\theta \circ_t \varepsilon_{t-3}\right) + \varepsilon_t, \tag{29}$$

where $\{\varepsilon_t\}$ is a sequence of i.i.d. Poisson random variables with parameter λ and $\theta \in (0, 1)$. The random vectors $(D_{t,1}, D_{t,2}, D_{t,3})$ are multinomial distribution with parameters $(1; \phi_1, \phi_2, \phi_3)$, independent of $\{\varepsilon_t\}$. The parameters values considered are

Table 1: Sample mean and mean square error (in brackets) for models A, B, and C.

Model	Parameter	Sample size 50	100	400	700
A	ϕ_1	0.0781 (0.0316)	0.0873 (0.0084)	0.0925 (0.0047)	0.0953 (0.0017)
	ϕ_2	0.1658 (0.0251)	0.1795 (0.0193)	0.1867 (0.0052)	0.1957 (0.0035)
	θ	0.2433 (0.0236)	0.2731 (0.0074)	0.2875 (0.0032)	0.2924 (0.0021)
	λ	1.1372 (0.1256)	1.1007 (0.0632)	1.0285 (0.0358)	1.0046 (0.0067)
B	ϕ_1	0.2547 (0.0412)	0.2871 (0.0136)	0.2935 (0.0087)	0.2977 (0.0032)
	ϕ_2	0.3468 (0.0308)	0.3726 (0.0158)	0.3891 (0.0072)	0.3964 (0.0023)
	θ	0.4625 (0.0217)	0.4789 (0.0104)	0.4836 (0.0044)	0.4976 (0.0013)
	λ	3.5672 (0.3265)	3.1741 (0.2381)	3.0522 (0.0943)	3.0075 (0.0045)
C	ϕ_1	0.5648 (0.0213)	0.5783 (0.0151)	0.5867 (0.0078)	0.5973 (0.0024)
	ϕ_2	0.1732 (0.0342)	0.1847 (0.0153)	0.1907 (0.0024)	0.1976 (0.0005)
	θ	0.6347 (0.0145)	0.6824 (0.0083)	0.6947 (0.0026)	0.6953 (0.0011)
	λ	12.3627 (1.0764)	10.6538 (0.3105)	10.1486 (0.0852)	10.0024 (0.0053)

(model A) $(\phi_1, \phi_2, \phi_3) = (0.1, 0.2, 0.7)$, $\theta = 0.3$, $\lambda = 1$,

(model B) $(\phi_1, \phi_2, \phi_3) = (0.3, 0.4, 0.3)$, $\theta = 0.5$, $\lambda = 3$,

(model C) $(\phi_1, \phi_2, \phi_3) = (0.6, 0.2, 0.2)$, $\theta = 0.7$, $\lambda = 10$.

The length of this discrete-valued time series T is 50, 100, 400, and 700. For each realization of these estimators, 500 independent replicates were simulated. The numerical results of the estimators for different true values of the parameters (ϕ_1, ϕ_2, ϕ_3), θ, and λ are presented in Table 1.

All the biases of ϕ_1, ϕ_2, and θ are negative, whereas the biases of λ are positive. It can be seen that as the sample size increases, the estimates seem to converge to the true parameter values. For example, when increasing sample size T, the bias and MSE both converge to zero. The reason might be that the Yule-Walker method is based on sufficient statistics. On the other hand, the performances of the estimators of λ are weaker than for the ones of ϕ_1 and θ.

5. Conclusion

In this paper, we introduce a class of self-exciting threshold integer-valued moving average models driven by decision random vector. Basic probabilistic and statistical properties of this class of models are discussed. Specifically, the method of estimation under analysis is the Yule-Walker. Their performance is compared through a simulation study. Potential issues of future research include extending the results for general INARMA(p, q) models including an arbitrary distribution of binomial thinning parameter as well as autoregressive and moving average parameters. This remains a topic of future research.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors are grateful to Professor Peter X.-K. Song from University of Michigan for his help and suggestions. The authors also thank an anonymous referee for helpful suggestions. This work is supported by the National Natural Science Foundation of China (no. 71171166 and no. 71201126), the Science Foundation of Ministry of Education of China (no. 12XJC910001), and the Fundamental Research Funds for the Central Universities (nos. JBK150949, JBK150501, and JBK120405).

References

[1] S. L. Zeger, "A regression model for time series of counts," *Biometrika*, vol. 75, no. 4, pp. 621–629, 1988.

[2] K. Brännäs and P. Johansson, "Time series count data regression," *Communications in Statistics: Theory and Methods*, vol. 23, no. 10, pp. 2907–2925, 1994.

[3] M. Al-Osh and A. A. Alzaid, "Integer-valued moving average (INMA) process," *Statistische Papers*, vol. 29, no. 4, pp. 281–300, 1988.

[4] R. Ferland, A. Latour, and D. Oraichi, "Integer-valued GARCH process," *Journal of Time Series Analysis*, vol. 27, no. 6, pp. 923–942, 2006.

[5] N. Silva, I. Pereira, and M. E. Silva, "Forecasting in INAR(1) model," *REVSTAT Statistical Journal*, vol. 7, no. 1, pp. 119–134, 2009.

[6] H. Zheng, I. V. Basawa, and S. Datta, "Inference for pth-order random coefficient integer-valued autoregressive processes," *Journal of Time Series Analysis*, vol. 27, no. 3, pp. 411–440, 2006.

[7] H. Zheng, I. V. Basawa, and S. Datta, "First-order random coefficient integer-valued autoregressive processes," *Journal of Statistical Planning and Inference*, vol. 137, no. 1, pp. 212–229, 2007.

[8] M. Kachour and L. Truquet, "A p-Order signed integer-valued autoregressive (SINAR(p)) model," *Journal of Time Series Analysis*, vol. 32, no. 3, pp. 223–236, 2011.

[9] M. M. Ristić, A. S. Nastić, and A. V. Miletić Ilić, "A geometric time series model with dependent Bernoulli counting series," *Journal of Time Series Analysis*, vol. 34, no. 4, pp. 466–476, 2013.

[10] C. H. Weiß, "The combined INAR (p) models for time series of counts," *Statistics & Probability Letters*, vol. 78, no. 13, pp. 1817–1822, 2008.

[11] M. M. Ristić, A. S. Nastić, and H. S. Bakouch, "Estimation in an integer-valued autoregressive process with negative binomial marginals (NBINAR(1))," *Communications in Statistics. Theory and Methods*, vol. 41, no. 4, pp. 606–618, 2012.

[12] S. Karlin and H. E. Taylor, *A First Course in Stochastic Processes*, Academic Press, New York, NY, USA, 1975.

[13] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*, Springer, New York, NY, USA, 1991.

[14] P. Billingsley, "Statistical methods in Markov chains," *Annals of Mathematical Statistics*, vol. 32, pp. 12–40, 1961.