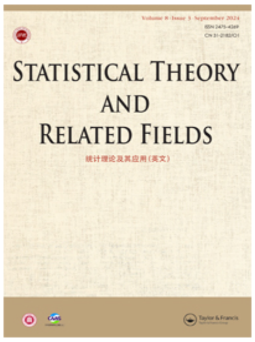

Statistical Theory and Related Fields

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tstf20

Maximum-likelihood estimation of the Po-MDDRCINAR(p) model with analysis of a COVID-19 data

Xiufang Liu, Jianlong Peng, Dehui Wang & Huaping Chen

To cite this article: Xiufang Liu, Jianlong Peng, Dehui Wang & Huaping Chen (09 Nov 2024): Maximum-likelihood estimation of the Po-MDDRCINAR(*p*) model with analysis of a COVID-19 data, Statistical Theory and Related Fields, DOI: 10.1080/24754269.2024.2412491

To link to this article: https://doi.org/10.1080/24754269.2024.2412491

Published online: 09 Nov 2024.

STATISTICAL THEORY AND RELATED FIELDS
https://doi.org/10.1080/24754269.2024.2412491

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

Maximum-likelihood estimation of the Po-MDDRCINAR(p) model with analysis of a COVID-19 data

Xiufang Liu [a], Jianlong Peng[a], Dehui Wang[b] and Huaping Chen[c]

[a]School of Mathematics, Taiyuan University of Technology, Taiyuan, People's Republic of China; [b]School of Mathematics and Statistics, Liaoning University, Shenyang, People's Republic of China; [c]School of Mathematics and Statistics, Henan University, Kaifeng, People's Republic of China

ABSTRACT

Integer-valued data are frequently encountered in time series studies. A pth-order mixed dependence-driven random coefficient integer-valued autoregressive time series model (Po-MDDRCINAR(p)) in view of binomial and negative binomial operators, where the innovation sequence follows a Poisson distribution, is investigated to provide meaningful theoretical explanations. Strict stationary and ergodicity of the model are demonstrated. Furthermore, the conditional least-squares and conditional maximum-likelihood methods are adopted to estimate the parameters, where the asymptotic characterization of the estimators is derived. Finite-sample properties of the conditional maximum-likelihood estimator are examined in relation to the widely used conditional least-squares estimator. The conclusion is that, if the Poisson assumption of the innovation sequence can be justified, conditional maximum-likelihood method performs better in terms of MADE and MSE. Finally, the practical performance of the model is illustrated by a set of COVID-19 data of suspected cases in China with a comparison with relevant models that exist so far in the literature.

ARTICLE HISTORY
Received 21 September 2023
Revised 16 July 2024
Accepted 24 September 2024

KEYWORDS
Po-MDDRCINAR(p) model; conditional least squares; conditional maximum-likelihood; asymptotic distribution; COVID-19 data

1. Introduction

In real-life situations, integer-valued time series data are very common. It is the integer-valued counting data formed by the state of a certain statistical index of a certain phenomenon at different times. This type of data is widely used in communication security, health care, law, actuarial insurance and many other fields, such as the number of patients admitted to a hospital in a region for a specific disease every month, the number of accidents in a region per month, etc. One of the most straightforward ways to construct integer-valued time series models is to use an appropriate thinning operator instead of the multiplication in the ARMA models. The first binomial thinning operator '$\circ$' was proposed by Steutel and Van Harn (1979), which is quite significant and prime to progress the integer-valued autoregressive models. The first-order integer-valued autoregressive (INAR(1)) process was proposed by Al-Osh and Alzaid (1987). It is defined as follows

$$X_t = \phi \circ X_{t-1} + \varepsilon_t, \quad t \geq 1,$$

where

$$\phi \circ X_{t-1} = \sum_{i=1}^{X_{t-1}} Y_i^{(t)}.$$

Here, the counting series $\{Y_i^{(t)}\}$ are independent and identically distributed (i.i.d.) Bernoulli random variables with success probability $\phi \in [0, 1]$ and $\{\varepsilon_t\}$ is a sequence of i.i.d. non-negative integer-valued random variables and independent of the counting series $\{Y_i^{(t)}\}$. Thus, $\phi \circ X_{t-1}$ is a binomial random variable with ϕ and X_{t-1} as parameters, namely, $\phi \circ X_{t-1} \sim B(X_{t-1}, \phi)$. This model has some similar characteristics to the ordinary AR(1) model, which has been discussed by Al-Osh and Aly (1992), Alzaid and Al-Osh (1988), Alzaid and Al-Osh (1990), among others.

Particularly, we focus on the fact that the model is related to a branching process with immigration having a Bernoulli offspring distribution. Here, an example of a standard INAR(1) model is presented, where X_t presents the number of surviving COVID-19 patients in a hospital at time t. Let ϕ be the probability of survival from time $t-1$ to t, and let ε_t denote the number of new COVID-19 patients admitted at time t. In fact, the parameter ϕ may vary with time and it may be random as the survival rate ϕ may be affected by various environmental factors, such as the

CONTACT Xiufang Liu liuxiufang@tyut.edu.cn Taiyuan University of Technology, No. 79 Yingze West Street, Taiyuan City, postal code 030024, Shanxi Province, People's Republic of China

quality of health care, the state of health of patients, etc. Therefore, it is indispensable to study the random coefficient INAR model. Recently, Zheng et al. (2006), Zheng et al. (2007) and Zheng and Basawa (2008) have researched the random coefficient integer-valued autoregressive model as well as Liu and Wang (2019), Liu et al. (2020) and Zhang and Wang (2015). In the above literature, the random coefficient INAR(p) model proposed by Zheng et al. (2006) is defined by the following equation:

$$X_t = \sum_{i=1}^{p} \alpha_i^{(t)} \circ X_{t-1} + \boldsymbol{\varepsilon}_t, \quad t \geq 1,$$

where $\alpha_i^{(t)}$ replaces the fixed α_i values in the literature by Du and Li (1991), and is an i.i.d sequence with a cumulative distribution function P_{α_i} on [0,1) with $E(\alpha_i^{(t)}) = \alpha_i$; $\{\varepsilon_t\}$ is a sequence of i.i.d. non-negative integer-valued random variables, independent of all the counting series and $\sum_{i=1}^{p} E(\alpha_i^{(t)}) < 1$;. Subsequently, the combined INAR(p) model was introduced by Weiß (2008). After that, a DDRCIN-AR(p) model was proposed by Liu and Wang (2019) to deal with the dependent data, as follows:

$$X_t = \sum_{i=1}^{p} \phi_{ti} \circ X_{t-1} + \boldsymbol{\varepsilon}_t, \quad t \geq 1,$$

where $\{\phi_{ti}\}$ is a dependence-driven sequence of random vectors with a joint distribution function $P_{\{\phi_{t1},\ldots,\phi_{tp}\}}$ given by

$$\begin{cases} p(\phi_{t1} = \phi_1, \phi_{t2} = 0, \ldots, \phi_{tp} = 0) = \alpha_1; \\ p(\phi_{t1} = 0, \phi_{t2} = \phi_2, \ldots, \phi_{tp} = 0) = \alpha_2; \\ \qquad\qquad \vdots \\ p(\phi_{t1} = 0, \phi_{t2} = 0, \ldots, \phi_{tp} = \phi_p) = \alpha_p; \\ p(\phi_{t1} = 0, \phi_{t2} = 0, \ldots, \phi_{tp} = 0) = \alpha_0. \end{cases}$$

Furthermore, Liu et al. (2020) studied the strictly stationary, ergodicity and estimation methods of parameters in the PoDDRCINAR(p) model.

In addition, Ristić et al. (2009) introduced a new first-order integer-valued autoregressive (NGINAR(1)) process, which has geometric marginals based on the negative binomial thinning operator, which contains the following geometric counting series

$$X_n = \alpha * X_{n-1} + \varepsilon_n, \quad n \geq 1,$$

where the operator '$*$' is defined as $\alpha * X = \sum_{i=1}^{X} W_i$, $\alpha \in [0, 1)$, X_n is a stationary process with Geometric $(\mu/(1+\mu))$ marginals, i.e. with probability mass function given by $P(X_n = x) = \mu^x/(1+\mu)^{x+1}$, $x = 0, 1, 2, \ldots$, $\{W_i\}$ is a sequence of independent identically distributed (i.i.d.) random variables with Geometric($\alpha/(1+\alpha)$) distribution, $\{\varepsilon_n\}$ is a sequence of i.i.d. random variables independent of $\{W_i\}$ and X_{n-l} and ε_n are independent for all $l \geq 1$. Furthermore, based on the negative binomial thinning operator, Nastić et al. (2012) proposed a combined Geometric INAR(p) model and Yu et al. (2019) studied a class of observation-driven random coefficient INAR(1) processes. Especially, if we assume that the ability of a random event to immediately cause one or more other events is relatively small, but this ability becomes more significant after a period of time, then both of the thinning operators mentioned above should be combined. In this case, Ristić and Nastić (2012) proposed a mixed INAR(p) model that is defined by the following equation:

$$X_t = \begin{cases} \alpha \circ_t X_{t-1} + \varepsilon_t, & \text{w.p. } \phi_1, \\ \alpha *_t X_{t-2} + \varepsilon_t, & \text{w.p. } \phi_2, \\ \alpha *_t X_{t-3} + \varepsilon_t, & \text{w.p. } \phi_3, \\ \qquad \vdots \\ \alpha *_t X_{t-p} + \varepsilon_t, & \text{w.p. } \phi_p, \end{cases}$$

where 'w.p.' means with probability, the binomial thinning operator is given by $\alpha \circ_t X = \sum_{i=1}^{X} Y_i^{(t)}$, where $\{Y_i^{(t)}\}$ is an i.i.d. sequence of Bernoulli(α) with $\alpha \in (0, 1)$, and the negative binomial thinning is given by $\alpha *_t X = \sum_{j=1}^{X} W_j^{(t)}$, where the count series $\{W_j^{(t)}\}$ is an i.i.d. sequence of Geometric($\alpha/(1+\alpha)$) and $\sum_{i=1}^{p} \phi_i = 1$. As analysed above, the parameter α of the mixed INAR(p) model by Ristić and Nastić (2012) may vary with time and may be random. Thus, Liu et al. (2023) proposed a MDDRCINAR(p) model, where $\{\alpha_{ti}\}$ is a dependence-driven sequence of random vectors with a similar joint distribution function to that of $\{\phi_{ti}\}$ in Liu and Wang (2019). According to the MDDRCINAR(p) model by Liu et al. (2023), this article mainly aims to study the basic statistical properties of this model with a Poisson innovation process and to introduce new inferential methods for the estimation of relevant parameters associated with the model.

The remainder of this paper is organized as follows. In Section 2, we introduce a Poisson mixed dependence-driven random coefficient integer-valued autoregressive (Po-MDDRCINAR(p)) model in detail, and we show, under certain conditions, stationarity and ergodicity of the Po-MDDRCINAR(p) model. In Section 3, we propose estimation methods for the parameters in the proposed model respectively and study their consistency and asymptotic properties. Furthermore, in Section 4, we conduct simulation studies to examine the finite-sample performance of the proposed procedures (Po-MDDRCINAR(2)). The proposed methods are illustrated with some real data sets in Section 5. In Section 6, we give a summary and concluding remarks. The paper ends with conditional moments used later, which are provided in Appendix.

2. The *p*th-order Poisson mixed dependence-driven random coefficient integer-valued autoregressive model

A pth-order Poisson mixed dependence-driven random coefficient integer-valued autoregressive (Po-MDDRCINAR(p)) model is defined by the following equation:

$$X_t = \alpha_{t1} \circ X_{t-1} + \alpha_{t2} * X_{t-2} + \cdots + \alpha_{tp} * X_{t-p} + \varepsilon_t, \quad t \geq 1, \tag{1}$$

with the joint distribution of $\{\alpha_{t1}, \alpha_{t2}, \ldots, \alpha_{tp}\}$ given by

$$\begin{cases} p(\alpha_{t1} = \alpha_1, \alpha_{t2} = 0, \ldots, \alpha_{tp} = 0) = \phi_1; \\ p(\alpha_{t1} = 0, \alpha_{t2} = \alpha_2, \ldots, \alpha_{tp} = 0) = \phi_2; \\ \qquad \vdots \\ p(\alpha_{t1} = 0, \alpha_{t2} = 0, \ldots, \alpha_{tp} = \alpha_p) = \phi_p; \\ p(\alpha_{t1} = 0, \alpha_{t2} = 0, \ldots, \alpha_{tp} = 0) = \phi_0, \end{cases} \tag{2}$$

where $\phi_0, \phi_1, \ldots, \phi_p$ are non-negative and $\sum_{i=0}^{p} \phi_i = 1$; $\{\varepsilon_t\}$ is an i.i.d. non-negative integer-valued sequence with a Poisson distribution Po(λ); the time-varying coefficients $\{\alpha_{ti}, 1 \leq i \leq p\}$ are i.i.d. random variables across the time points t; $\{\alpha_{ti}, 1 \leq i \leq p\}$ and $\{\varepsilon_t\}$ are independent sequences; the counting series in $\alpha_{t1} \circ X_{t-1}$ and $\alpha_{ti} * X_{t-i} (i = 2, \ldots, p)$ are named as survival processes and are mutually independent for all $t \in Z$ for known X_{t-i}.

The model defined by Equations (1) and (2) can be formulated by the following form

$$X_t = \begin{cases} \alpha_1 \circ X_{t-1} & \text{w.p.} & \phi_1 \\ \alpha_2 * X_{t-2} & \text{w.p.} & \phi_2 \\ \vdots & \vdots & \vdots \\ \alpha_p * X_{t-p} & \text{w.p.} & \phi_p \\ 0 & \text{w.p.} & \phi_0 \end{cases} + \varepsilon_t. \tag{3}$$

Therefore, we can derive

$$E(\varepsilon_t) = \text{Var}(\varepsilon_t) = \lambda, \tag{4}$$

$$E(\alpha_{ti}) = \phi_i \alpha_i, \quad \text{Var}(\alpha_{ti}) = \phi_i(1 - \phi_i)\alpha_i^2, \quad \text{Cov}(\alpha_{ti}, \alpha_{tj}) = -\phi_i \phi_j \alpha_i \alpha_j \ (i \neq j). \tag{5}$$

The model includes several special cases:

(i) For $\phi_2 = \cdots = \phi_p = \phi_0 = 0$, the model $\{X_t\}$ reduces to the first-order integer-valued autoregressive model (INAR(1)) by the binomial thinning operator in Al-Osh and Alzaid (1987).
(ii) For $\phi_1 = \cdots = \phi_p = 0$, the model $\{X_t\}$ degenerates to the innovation process $\{\varepsilon_t\}$.
(iii) For $\phi_0 = 0$ and $\alpha_1 = \alpha_2 = \cdots = \alpha_p = \alpha$, the model reduces to the mixed INAR(p) model proposed by Ristić and Nastić (2012).

Now, we are ready to show the high-order moments of $\{X_t\}$ and to discuss its stochastic properties. First, we introduce some notation.

Let

$$\begin{aligned}
&\boldsymbol{X} = (X_1, X_2, \dots, X_p)^\top, \quad \boldsymbol{Z} = (Z_1, Z_2, \dots, Z_p)^\top, \\
&\boldsymbol{X}_t = (X_t, X_{t-1}, \dots, X_{t-p+1})^\top_{p\times 1}, \quad \boldsymbol{\varepsilon}_t = (\varepsilon_t, 0, \dots, 0)^\top_{p\times 1}, \\
&\boldsymbol{A}_t = \begin{bmatrix} \alpha_1^{(t)} & \alpha_2^{(t)} & \dots & \alpha_{p-1}^{(t)} & \alpha_p^{(t)} \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}, \quad \boldsymbol{B}_t = \begin{bmatrix} \beta_1^{(t)} & \beta_2^{(t)} & \dots & \beta_{p-1}^{(t)} & \beta_p^{(t)} \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}, \\
&\boldsymbol{A} = \begin{bmatrix} \phi_1\alpha_1 & \phi_2\alpha_2 & \dots & \phi_{p-1}\alpha_{p-1} & \phi_p\alpha_p \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}, \\
&\boldsymbol{B} = \begin{bmatrix} \phi_1\beta_1 & \phi_2\beta_2 & \dots & \phi_{p-1}\beta_{p-1} & \phi_p\beta_p \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix},
\end{aligned}$$

where $\boldsymbol{A}_t$ and $\boldsymbol{B}_t$ are independent. We then use the notation

$$\boldsymbol{A}_t \circ \boldsymbol{X} = \left(\sum_{j=1}^{p} (\alpha_j^{(t)} \circ X_j), X_1, \dots, X_{p-1} \right)^\top,$$

where $\{\alpha_1^{(t)}, \alpha_2^{(t)}, \dots, \alpha_p^{(t)}\}$ also satisfy Equation (2). $\alpha_j^{(t)} \circ$ is a thinning operator based on Bernoulli($\alpha_1^{(t)}$) counting series if $j = 1$, and on Geometric($\alpha_j^{(t)}/(1 + \alpha_j^{(t)})$) counting series if $j = 2, \dots, p$. And

$$\boldsymbol{X}_t = \boldsymbol{A}_t \circ \boldsymbol{X}_{t-1} + \boldsymbol{\varepsilon}_t. \tag{6}$$

Next, the following assumption is significant in Propositions 2.1–2.3.

Assumption 2.1: *(i).* $\sum_{i=1}^{p} \phi_i\alpha_i < 1, \ \sum_{i=0}^{p} \phi_i = 1, \ 0 \le \phi_i \le 1$. *(ii). The parametric space* Θ *is compact with* $\Theta = \{\boldsymbol{\theta} \,|\, \boldsymbol{\theta} = (\alpha_1, \alpha_2, \dots, \alpha_p, \phi_1, \phi_2, \dots, \phi_p, \lambda)^\top, \underline{\alpha} \le \alpha_i \le \bar{\alpha}, \ \underline{\phi} \le \phi \le \bar{\phi}, \ (i = 1, 2, \dots, p), \underline{\lambda} \le \lambda \le \bar{\lambda}$ *and* $\underline{\gamma} \le \sum_{i=1}^{p} \phi_i\alpha_i \le \bar{\gamma} < 1\}$, *where* $\underline{\alpha}, \bar{\alpha}, \ \underline{\phi}, \bar{\phi}, \ \underline{\lambda}, \bar{\lambda}, \ \underline{\gamma}$ *and* $\bar{\gamma}$ *are finite positive constants, and* $\boldsymbol{\theta}_0$ *is an interior point in* Θ.

Proposition 2.1: *For* $t \ge 1$

(i) $E(\boldsymbol{A}_t \circ X \,|\, X) = \boldsymbol{AX}$, $E(\mathbf{A}_t \circ \boldsymbol{X}) = \boldsymbol{A}E(\boldsymbol{X})$;
(ii) $E[(\boldsymbol{A}_t \circ \boldsymbol{X})\boldsymbol{Z}^\top] = \boldsymbol{A}E(\boldsymbol{XZ}^\top)$;
(iii) $E[(\boldsymbol{A}_t \circ \boldsymbol{X})(\boldsymbol{B}_t \circ \boldsymbol{Z})^\top] = \boldsymbol{A}E(\boldsymbol{XZ}^\top)B^\top$, *if all the counting series of* $\boldsymbol{A}_t \circ \boldsymbol{X}$ *and* $\boldsymbol{B}_t \circ \boldsymbol{Z}$ *are independent;*
(iv) $E[(\boldsymbol{A}_t \circ \boldsymbol{X})(\boldsymbol{A}_t \circ \boldsymbol{X})^\top \,|\, \boldsymbol{X}, \mathcal{J}_{tp}] = \boldsymbol{A}_t\boldsymbol{XX}^\top\boldsymbol{A}_t^\top + \boldsymbol{D}_1$, $E[(\boldsymbol{A}_t \circ \boldsymbol{X})(\boldsymbol{A}_t \circ \boldsymbol{X})^\top \,|\, \boldsymbol{X}] = \boldsymbol{AXX}^\top\boldsymbol{A}^\top + \boldsymbol{D}_2$ *and* $E[(\boldsymbol{A}_t \circ \boldsymbol{X})(\boldsymbol{A}_t \circ \boldsymbol{X})^\top] = \boldsymbol{A}E(\boldsymbol{XX}^\top)\boldsymbol{A}^\top + \boldsymbol{D}_3$, *where*

$$\boldsymbol{D}_1 = \begin{bmatrix} \alpha_1^{(t)}(1 - \alpha_1^{(t)})X_1 + \sum_{i=2}^{p} \alpha_i^{(t)}(1 + \alpha_i^{(t)})X_i & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}_{p\times p},$$

$$\boldsymbol{D}_2 = \begin{bmatrix} \sum_{i=1}^{p} \phi_i\alpha_i^2X_i^2 + \phi_1\alpha_1(1 - \alpha_1)X_1 + \sum_{i=2}^{p} \phi_i\alpha_i(1 + \alpha_i)X_i - (\sum_{i=1}^{p} \phi_i\alpha_iX_i)^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}_{p\times p},$$

$$\boldsymbol{D}_3 = \begin{bmatrix} d_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}_{p\times p}$$

with

$$d_3 = \sum_{i=1}^{p} \phi_i(1-\phi_i)\alpha_i^2 E(X_i^2) + \phi_1\alpha_1(1-\alpha_1)E(X_1) + \sum_{i=2}^{p} \phi_i\alpha_i(1+\alpha_i)E(X_i) - \sum_{i=2}^{p}\sum_{j=1}^{i-1} \phi_i\phi_j\alpha_i\alpha_j E(X_iX_j)$$

and $\mathcal{J}_{tp}$ being the σ-field generated by $\alpha_1^{(t)}, \alpha_2^{(t)}, \ldots, \alpha_p^{(t)}$.

Proof: The proof is similar to the proof of Lemma 2.1 in Du and Li (1991), Lemma 2 in Ristić and Nastić (2012) and Proposition 2.1 in Liu et al. (2020). ■

Proposition 2.2: *Let $\{X_t\}$ be the Po-MDDRCINAR(p) model defined by Equations* (1) *and* (2)*. Then*

(i) $E(X_t \mid X_{t-1}, \ldots, X_{t-p}) = \sum_{i=1}^{p} \phi_i\alpha_i X_{t-i} + \lambda$;
(ii) $E(X_t) = \frac{\lambda}{1-\sum_{i=1}^{p}\phi_i\alpha_i}$, *if the model is first-order stationary;*
(iii) $\mathrm{Var}(X_t \mid X_{t-1}, \ldots, X_{t-p}, \mathcal{J}_t) = \alpha_{t1}(1-\alpha_{t1})X_{t-1} + \sum_{i=2}^{p} \alpha_{ti}(1+\alpha_{ti})X_{t-i} + \lambda$, *where $\mathcal{J}_t$ is the σ-field generated by $\alpha_{t1}, \alpha_{t2}, \ldots, \alpha_{tp}$;*
(iv) $\mathrm{Var}(X_t \mid X_{t-1}, \ldots, X_{t-p}) = \sum_{i=1}^{p} \phi_i(1-\phi_i)\alpha_i^2 X_{t-i}^2 + \phi_1\alpha_1(1-\alpha_1)X_{t-1} + \sum_{i=2}^{p} \phi_i\alpha_i(1+\alpha_i)X_{t-i} - 2\sum_{j=2}^{p}\sum_{i=1}^{j-1} \alpha_i\alpha_j\phi_i\phi_j X_{t-i}X_{t-j} + \lambda$;
(v) *Let $\gamma_k = \mathrm{Cov}(X_t, X_{t-k})$, and then $\gamma_k = \sum_{i=1}^{p} \phi_i\alpha_i\gamma_{k-i}$.*

Proof: The proofs for results (i)-(iv) are straightforward. The proof of (v) follows from Proposition 1–(v) in Zheng et al. (2006). ■

Then we prove the strict stationary and ergodicity of Po-MDDRCINAR(p) process defined by Equations (1) and (2). These properties will be useful to derive the asymptotic properties of parameter estimators in the model.

Proposition 2.3: *The Po-MDDRCINAR(p) model $\{X_t\}$ defined by Equations* (1) *and* (2) *is strictly stationary and ergodic.*

Proof: According to the proof of Theorem 2 of mixed INAR(p) model proposed by Ristić and Nastić (2012), the Proposition can be obtained. ■

Next, we consider the problem of estimation involved in the Po-MDDRCINAR(p) model.

3. Estimation methods

3.1. Conditional least squares (CLS) estimators

We re-parameterize Equation (5) by defining

$$\beta_i = \alpha_i\phi_i, \quad \sigma_{ii} = \phi_i\alpha_i^2(1-\phi_i),\ i = 1, 2, \ldots, p. \tag{7}$$

Let $\mathcal{F}_{t-1}$ be the σ-field generated by $X_1, X_2, \ldots, X_{t-1}$. Denote

$$\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)^\top, \quad \boldsymbol{\sigma} = (\sigma_{11}, \sigma_{22}, \ldots, \sigma_{pp})^\top,$$
$$\boldsymbol{\vartheta} = (\boldsymbol{\beta}', \boldsymbol{\sigma}', \lambda)^\top, \quad \boldsymbol{\theta} = (\alpha_1, \alpha_2, \ldots, \alpha_p, \phi_1, \phi_2, \ldots, \phi_p, \lambda)^\top.$$

Assume that observations of X_t are available for $t = 1, 2, \ldots, n$.

The CLS estimates of $\boldsymbol{\beta}$ and λ can be obtained by minimizing

$$Q_1(\boldsymbol{\beta},\lambda) = \sum_{t=p+1}^{n} u_t^2 = \sum_{t=p+1}^{n}\left(X_t - \sum_{i=1}^{p}\phi_i\alpha_i X_{t-i} - \lambda\right)^2 = \sum_{t=p+1}^{n}(X_t - \boldsymbol{Y}_t^\top\boldsymbol{\beta} - \lambda)^2$$

with respect to $\boldsymbol{\beta}$ and λ, where $u_t = X_t - E(X_t \mid \mathcal{F}_{t-1})$. This yields the estimators

$$\hat{\boldsymbol{\beta}} = \left(\sum_{t=p+1}^{n}\boldsymbol{Y}_t\boldsymbol{Y}_t^\top - \frac{1}{n-p}\sum_{t=p+1}^{n}\boldsymbol{Y}_t\sum_{t=p+1}^{n}\boldsymbol{Y}_t^\top\right)^{-1} \times\left(\sum_{t=p+1}^{n}\boldsymbol{Y}_tX_t - \frac{1}{n-p}\sum_{t=p+1}^{n}\boldsymbol{Y}_t\sum_{t=p+1}^{n}X_t\right), \tag{8}$$

$$\hat{\lambda} = \frac{1}{n-p}\sum_{t=p+1}^{n}(X_t - \boldsymbol{Y}_t^\top\hat{\boldsymbol{\beta}}) \tag{9}$$

with $\boldsymbol{Y}_t = (X_{t-1}, X_{t-2}, \ldots, X_{t-p})^\top$.

To obtain estimates of $\boldsymbol{\sigma}$, CLS is again applied to estimate the residual sequence K_t, where

$$\widehat{K}_t = \left(X_t - \sum_{i=1}^{p}\hat{\beta}_iX_{t-i} - \hat{\lambda}\right)^2 + 2\sum_{j=2}^{p}\sum_{i=1}^{j-1}\hat{\beta}_i\hat{\beta}_jX_{t-i}X_{t-j} - (\hat{\beta}_1 - \hat{\beta}_1^2)X_{t-1} - \sum_{i=2}^{p}(\hat{\beta}_i + \hat{\beta}_i^2)X_{t-i} - \hat{\lambda},$$

Minimizing

$$Q_2(\boldsymbol{\sigma}) = \sum_{t=p+1}^{n}\left(\widehat{K}_t - \sigma_{11}(X_{t-1}^2 - X_{t-1}) - \sum_{i=2}^{p}\sigma_{ii}(X_{t-i}^2 + X_{t-i})\right)^2 = \sum_{t=p+1}^{n}(\widehat{K}_t - \boldsymbol{Z}_t^\top\boldsymbol{\sigma})^2$$

with respect to $\boldsymbol{\sigma}$ yields the estimators of $\boldsymbol{\sigma}$

$$\hat{\boldsymbol{\sigma}} = \left(\sum_{t=p+1}^{n}\boldsymbol{Z}_t\boldsymbol{Z}_t^\top\right)^{-1}\sum_{t=p+1}^{n}\boldsymbol{Z}_t\widehat{K}_t, \tag{10}$$

where $\boldsymbol{Z}_t = (X_{t-1}^2 - X_{t-1}, X_{t-2}^2 + X_{t-2}, \ldots, X_{t-p}^2 + X_{t-p})^\top$.

Now, the estimates $\hat{\boldsymbol{\theta}}_{CLS}$ can be obtained as

$$\hat{\phi}_i = \frac{\hat{\beta}_i^2}{\hat{\sigma}_{ii} + \hat{\beta}_i^2} \text{ and } \hat{\alpha}_i = \frac{\hat{\sigma}_{ii} + \hat{\beta}_i^2}{\hat{\beta}_i}, \quad i = 1, 2, \ldots, p. \tag{11}$$

The following theorem gives the strong consistency and the limited distribution for the estimates $\hat{\boldsymbol{\theta}}_{CLS}$ given in Equation (11).

Theorem 3.1: *Let $\{X_t\}$ be an Po-MDDRCINAR(p) process generated as in Equations* (1) *and* (2). *Then the estimates $\hat{\boldsymbol{\theta}}_{CLS}$ obtained from Equation* (11) *will be strongly consistent and jointly asymptotically normally distributed.*

Proof: First, the strong consistency of $\hat{\boldsymbol{\theta}}_{CLS}$ can be easily obtained according to Theorem 2.1 of Klimko and Nelson (1978), the proof of which is omitted here.

Next, based on the Assumption 2.1 and Proposition 2.3, we can prove the asymptotic normality of the estimates $\hat{\boldsymbol{\theta}}_{CLS}$. According to Theorem 3.1 of Hwang and Basawa (1998) or Theorem 3.1 of Nicholls and Quinn (1982), we have

$$\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{d} N_p(\boldsymbol{0}, \boldsymbol{L}^{-1}\boldsymbol{W}\boldsymbol{L}^{-1}),$$

where $\boldsymbol{W} = E(u_t^2\boldsymbol{Y}_t\boldsymbol{Y}_t^\top) = E(\mathrm{Var}(X_t \mid \mathcal{F}_{t-1})\boldsymbol{Y}_t\boldsymbol{Y}_t^\top)$ and $\boldsymbol{L} = E(\boldsymbol{Y}_t\boldsymbol{Y}_t^\top)$. With the similar method, we can obtain

$$\sqrt{n}(\hat{\boldsymbol{\sigma}} - \boldsymbol{\sigma}) \xrightarrow{d} N_p(\boldsymbol{0}, \boldsymbol{\Gamma}^{-1}\boldsymbol{\Sigma}\boldsymbol{\Gamma}^{-1}),$$

$$\sqrt{n}(\hat{\lambda} - \lambda) \xrightarrow{d} N(0, I).$$

According to Theorem 3.2 in Nicholls and Quinn (1982), we have

$$\sqrt{n}(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}) \xrightarrow{d} N_{2p+1}(\mathbf{0}, \boldsymbol{\Theta}),$$

where

$$\boldsymbol{\Theta} = \begin{bmatrix} \boldsymbol{L}^{-1}\mathbf{W}\boldsymbol{L}^{-1} & \boldsymbol{L}^{-1}\boldsymbol{M}\boldsymbol{\Gamma}^{-1} & \boldsymbol{L}^{-1}\boldsymbol{V} \\ \boldsymbol{\Gamma}^{-1}\boldsymbol{M}^{\top}\boldsymbol{L}^{-1} & \boldsymbol{\Gamma}^{-1}\boldsymbol{\Sigma}\boldsymbol{\Gamma}^{-1} & \boldsymbol{\Gamma}^{-1}\boldsymbol{D} \\ \boldsymbol{V}^{\top}\boldsymbol{L}^{-1} & \boldsymbol{D}^{\top}\boldsymbol{\Gamma}^{-1} & \boldsymbol{I} \end{bmatrix},$$

and

$$\boldsymbol{V} = E(\boldsymbol{Y}_t u_t^2), \quad \boldsymbol{M} = E(u_t U_t \boldsymbol{Y}_t \boldsymbol{Z}_t^{\top}) = E(\varphi_t \boldsymbol{Y}_t \boldsymbol{Z}_t^{\top}), \quad \boldsymbol{\Gamma} = E(\boldsymbol{Z}_t \boldsymbol{Z}_t^{\top}),$$
$$\boldsymbol{I} = E(u_t^2), \quad \boldsymbol{D} = E(\boldsymbol{Z}_t U_t u_t), \quad \boldsymbol{\Sigma} = E(U_t^2 \boldsymbol{Z}_t \boldsymbol{Z}_t^{\top}),$$

where

$$\varphi_t = E(X_t^3 \mid \mathcal{F}_{t-1}) - 3\mathrm{Var}(X_t \mid \mathcal{F}_{t-1})E(X_t \mid \mathcal{F}_{t-1}) - (E(X_t \mid \mathcal{F}_{t-1}))^3,$$
$$U_t = u_t^2 - E(u_t^2 \mid \mathcal{F}_{t-1}).$$

According to Equation (11) and Proposition 6.4.3 of Brockwell and Davis (1987), we have

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_{CLS} - \boldsymbol{\theta}) \xrightarrow{d} N_{2p+1}(\mathbf{0}, \boldsymbol{\Omega}\boldsymbol{\Theta}\boldsymbol{\Omega}^{\top}),$$

where

$$\boldsymbol{\Omega} = \begin{bmatrix} \boldsymbol{\Omega}_{11} & \boldsymbol{\Omega}_{12} & \boldsymbol{\Omega}_{13} \\ \boldsymbol{\Omega}_{21} & \boldsymbol{\Omega}_{22} & \boldsymbol{\Omega}_{23} \\ \boldsymbol{\Omega}_{31} & \boldsymbol{\Omega}_{32} & \boldsymbol{\Omega}_{33} \end{bmatrix}$$

with

$$\boldsymbol{\Omega}_{11} = \mathrm{diag}\left(\frac{\beta_1^2 - \sigma_{11}}{\beta_1^2}, \frac{\beta_2^2 - \sigma_{22}}{\beta_2^2}, \ldots, \frac{\beta_p^2 - \sigma_{pp}}{\beta_p^2}\right),$$
$$\boldsymbol{\Omega}_{12} = \mathrm{diag}\left(\frac{1}{\beta_1}, \frac{1}{\beta_2}, \ldots, \frac{1}{\beta_p}\right),$$
$$\boldsymbol{\Omega}_{13} = \boldsymbol{\Omega}_{23} = \mathbf{0}_{p\times 1},$$
$$\boldsymbol{\Omega}_{21} = \mathrm{diag}\left(\frac{2\beta_1\sigma_{11}}{(\sigma_{11} + \beta_1^2)^2}, \frac{2\beta_2\sigma_{22}}{(\sigma_{22} + \beta_2^2)^2}, \ldots, \frac{2\beta_p\sigma_{pp}}{(\sigma_{pp} + \beta_p^2)^2}\right),$$
$$\boldsymbol{\Omega}_{22} = \mathrm{diag}\left(\frac{-\beta_1^2}{(\sigma_{11} + \beta_1^2)^2}, \frac{-\beta_2^2}{(\sigma_{22} + \beta_2^2)^2}, \ldots, \frac{-\beta_p^2}{(\sigma_{pp} + \beta_p^2)^2}\right),$$
$$\boldsymbol{\Omega}_{31} = \boldsymbol{\Omega}_{32} = \mathbf{0}_{1\times p},$$
$$\boldsymbol{\Omega}_{33} = 1.$$

■

Remark 3.1: The detailed proof of the strong consistency of $\hat{\boldsymbol{\theta}}_{CLS}$ can refer to the proof process of Theorem 3.1 in Liu et al. (2023).

3.2. Conditional maximum-likelihood estimators

We now consider the conditional maximum-likelihood (CML) estimators of parameters for model Po-MDDRCINAR(p) defined by Equations (1) and (2), or by Equation (3). Suppose that $X_0, X_1, \ldots, X_n$ are generated by the Po-MDDRCINAR(p) model with the true parameter value θ_0.

Then, the conditional log-likelihood function for sample observation $X_1, \ldots, X_{n+p+1}$ can be written as

$$
\begin{aligned}
\ell(\boldsymbol{\theta}) &= \sum_{t=p+1}^{n} \log P_{\boldsymbol{\theta}}(X_t \mid X_{t-1}, \ldots, X_{t-p}) \\
&= \sum_{t=p+1}^{n} \log \Bigg[\phi_1 \sum_{k_1=0}^{\min(X_t, X_{t-1})} \tbinom{X_{t-1}}{k_1} \alpha_1^{k_1} (1-\alpha_1)^{X_{t-1}-k_1} \frac{\lambda^{X_t - k_1} \mathrm{e}^{-\lambda}}{(X_t - k_1)!} \\
&\quad + \sum_{i=2}^{p} \phi_i \sum_{k_i=0}^{X_t} \tbinom{X_{t-i}+k_i-1}{X_{t-i}-1} \left(\frac{\alpha_i}{1+\alpha_i}\right)^{k_i} \left(\frac{1}{1+\alpha_i}\right)^{X_{t-i}} \frac{\lambda^{X_t - k_i} \mathrm{e}^{-\lambda}}{(X_t - k_i)!} \\
&\quad + \left(1 - \sum_{i=1}^{p} \phi_i\right) \frac{\lambda^{X_t} \mathrm{e}^{-\lambda}}{X_t!} \Bigg] \\
&= \sum_{t=p+1}^{n} \log \Bigg(\phi_1 \sum_{k_1=0}^{\min(X_t, X_{t-1})} Z(X_{t-1}, k_1) q(X_t, k_1) \\
&\quad + \sum_{i=2}^{p} \phi_i \sum_{k_i=0}^{X_t} Y(X_{t-i}, k_i) q(X_t, k_i) + \left(1 - \sum_{i=1}^{p} \phi_i\right) q(X_t, k_0) \Bigg),
\end{aligned}
$$

where

$$
\begin{aligned}
Z(X_{t-1}, k_1) &= \binom{X_{t-1}}{k_1} \alpha_1^{k_1} (1-\alpha_1)^{X_{t-1}-k_1}, \\
Y(X_{t-i}, k_i) &= \binom{X_{t-i}+k_i-1}{X_{t-i}-1} \left(\frac{\alpha_i}{1+\alpha_i}\right)^{k_i} \left(\frac{1}{1+\alpha_i}\right)^{X_{t-i}} \quad (i = 2, \ldots, p), \\
q(X_t, k_i) &= \frac{\lambda^{X_t - k_i} \mathrm{e}^{-\lambda}}{(X_t - k_i)!} \quad (i = 0, 1, \ldots, p)
\end{aligned}
$$

with $k_0 = 0$.

Thus, the score function for the model is presented by

$$
\begin{aligned}
\frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha_1} &= \sum_{t=p+1}^{n} \frac{1}{P_{\boldsymbol{\theta}}(X_t \mid X_{t-1}, \ldots, X_{t-p})} \\
&\quad \times \Bigg[\frac{\phi_1 X_{t-1}}{1-\alpha_1} \Bigg(\sum_{k_1=0}^{\min(X_t-1, X_{t-1}-1)} Z(X_{t-1}-1, k_1) q(X_t - 1, k_1) \\
&\quad - \sum_{k_1=0}^{\min(X_t, X_{t-1})} Z(X_{t-1}, k_1) q(X_t, k_1) \Bigg) \Bigg], \\
\frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha_i} &= \sum_{t=p+1}^{n} \frac{1}{P_{\boldsymbol{\theta}}(X_t \mid X_{t-1}, \ldots, X_{t-p})} \Bigg[\frac{\phi_i X_{t-i}}{1+\alpha_i} \Bigg(\sum_{k_i=0}^{X_t - 1} Y(X_{t-i}+1, k_i) q(X_t - 1, k_i) \\
&\quad - \sum_{k_i=0}^{X_t} Y(X_{t-i}, k_i) q(X_t, k_i) \Bigg) \Bigg], \quad (i = 2, \ldots, p) \\
\frac{\partial \ell(\boldsymbol{\theta})}{\partial \phi_1} &= \sum_{t=p+1}^{n} \frac{1}{P_{\boldsymbol{\theta}}(X_t \mid X_{t-1}, \ldots, X_{t-p})} \\
&\quad \times \left(\sum_{k_1=0}^{\min(X_t, X_{t-1})} Z(X_{t-1}, k_1) q(X_t, k_1) - q(X_t, k_0) \right),
\end{aligned}
$$

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \phi_i} = \sum_{t=p+1}^{n} \frac{1}{P_{\boldsymbol{\theta}}(X_t \mid X_{t-1},\ldots,X_{t-p})} \left(\sum_{k_i=0}^{X_t} Y(X_{t-i},k_i)q(X_t,k_i) - q(X_t,k_0) \right), \quad (i=2,\ldots,p)$$

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \lambda} = \sum_{t=p+1}^{n} \frac{1}{P_{\boldsymbol{\theta}}(X_t \mid X_{t-1},\ldots,X_{t-p})} \Bigg[\phi_1 \sum_{k_1=0}^{\min(X_t-1,X_{t-1})} Z(X_{t-1},k_1)q(X_t-1,k_1)$$
$$+ \sum_{i=2}^{p} \phi_i \sum_{k_i=0}^{X_t-1} Y(X_{t-i},k_i)q(X_t-1,k_i) + \left(1-\sum_{i=1}^{p}\phi_i\right) q(X_t-1,k_0)$$
$$-P_{\boldsymbol{\theta}}(X_t \mid X_{t-1},\ldots,X_{t-p}) \Bigg],$$

The CML estimators of the unknown parameters $\boldsymbol{\theta} = (\alpha_1,\ldots,\alpha_p,\phi_1,\ldots,\phi_p,\lambda)^\top$ are derived numerically by maximizing the conditional log-likelihood function or by solving the equations related to the score functions.

Next, we consider the asymptotic behaviours of CML estimates for Po-MDDRCINAR(p) parameters. At first, the transition probability $P_{\boldsymbol{\theta}}(X_t \mid X_{t-1},\ldots,X_{t-p})$ can be written as

$$P(X_t,X_{t-1},\ldots,X_{t-p}) := P_{\boldsymbol{\theta}}(X_t \mid X_{t-1},\ldots,X_{t-p})$$
$$= \phi_1 \sum_{k_1=0}^{\min(X_t,X_{t-1})} Z(X_{t-1},k_1)q(X_t,k_1)$$
$$+ \sum_{i=2}^{p} \phi_i \sum_{k_i=0}^{X_t} Y(X_{t-i},k_i)q(X_t,k_i)$$
$$+ \left(1-\sum_{i=1}^{p}\phi_i\right) q(X_t,k_0)$$

with $X_i(i=t,t-1,\ldots,t-p) \geq 0$.

The following lemma shows the useful recurrence relation for Z, Y, P and their partial derivatives.

Lemma 3.2: (a) $Z(X_{t-1},k_1) = \alpha_1 Z(X_{t-1}-1,k_1-1) + (1-\alpha_1)Z(X_{t-1}-1,k_1)$.
(b) $Y(X_{t-i},k_i) = \frac{\alpha_i X_{t-i}}{k_i} Y(X_{t-i}+1,k_i-1)$, $i=2,\ldots,p$.
(c) $\frac{\partial Z(X_{t-1},k_1)}{\partial \alpha_1} = \frac{X_{t-1}}{1-\alpha_1}[Z(X_{t-1}-1,k_1-1) - Z(X_{t-1},k_1)]$.
(d) $\frac{\partial Y(X_{t-i},k_i)}{\partial \alpha_i} = \frac{X_{t-i}}{1+\alpha_i}[Y(X_{t-i}+1,k_i-1) - Y(X_{t-i},k_i)]$, $i=2,\ldots,p$.
(e) $\frac{\partial P(X_t,X_{t-1},\ldots,X_{t-p})}{\partial \alpha_1} = \phi_1 \sum_{k_1=0}^{\min(X_t,X_{t-1})} \frac{\partial Z(X_{t-1},k_1)}{\partial \alpha_1} q(X_t,k_1)$.
(f) $\frac{\partial P(X_t,X_{t-1},\ldots,X_{t-p})}{\partial \alpha_i} = \phi_i \sum_{k_i=0}^{X_t} \frac{\partial Y(X_{t-i},k_i)}{\partial \alpha_i} q(X_t,k_i)$, $i=2,\ldots,p$.

Proof: The proof is similar to the proof of Lemma 4 in Shirozhan and Mohammadpour (2020). ∎

The consequences of Lemma 3.1 are

$$-\frac{X_{t-1}}{1-\alpha_1} \leq \frac{\partial \log P(X_t,X_{t-1},\ldots,X_{t-p})}{\partial \alpha_1} \leq \frac{X_{t-1}}{\alpha_1(1-\alpha_1)} \tag{12}$$

and

$$-\frac{X_{t-i}}{1+\alpha_i} \leq \frac{\partial \log P(X_t,X_{t-1},\ldots,X_{t-p})}{\partial \alpha_i} \leq \frac{X_t}{\alpha_i(1+\alpha_i)}, \quad i=2,\ldots,p. \tag{13}$$

Since

$$\frac{\partial \log P(X_t,X_{t-1},\ldots,X_{t-p})}{\partial \alpha_1}$$
$$= \frac{\phi_1 X_{t-1}}{(1-\alpha_1)P(X_t,X_{t-1},\ldots,X_{t-p})}$$
$$\times \Bigg(\sum_{k_1=0}^{\min(X_t-1,X_{t-1}-1)} Z(X_{t-1}-1,k_1)q(X_t-1,k_1)$$

$$
\left. - \sum_{k_1=0}^{\min(X_t,X_{t-1})} Z(X_{t-1},k_1)q(X_t,k_1) \right)
$$

$$
\geq - \frac{\frac{\phi_1 X_{t-1}}{1-\alpha_1} \sum_{k_1=0}^{\min(X_t,X_{t-1})} Z(X_{t-1},k_1)q(X_t,k_1)}{P(X_t,X_{t-1},\ldots,X_{t-p})}
$$

$$
\geq - \frac{X_{t-1}}{1-\alpha_1},
$$

and on the other hand,

$$
\frac{\partial \log P(X_t,X_{t-1},\ldots,X_{t-p})}{\partial \alpha}
$$

$$
\leq \frac{\frac{\phi_1 X_{t-1}}{\alpha_1(1-\alpha_1)} \sum_{k_1=0}^{\min(X_t,X_{t-1})} (Z(X_{t-1},k_1) - (1-\alpha_1)Z(X_{t-1}-1,k_1))q(X_t,k_1)}{P(X_t,X_{t-1},\ldots,X_{t-p})}
$$

$$
\leq \frac{\frac{\phi_1 X_{t-1}}{\alpha_1(1-\alpha_1)} \sum_{k_1=0}^{\min(X_t,X_{t-1})} Z(X_{t-1},k_1)q(X_t,k_1)}{P(X_t,X_{t-1},\ldots,X_{t-p})}
$$

$$
\leq \frac{X_{t-1}}{\alpha_1(1-\alpha_1)}.
$$

Then,

$$
\frac{\partial \log P(X_t,X_{t-1},\ldots,X_{t-p})}{\partial \alpha_i}
$$

$$
= \frac{\frac{\phi_i X_{t-i}}{1+\alpha_i} \left(\sum_{k_i=0}^{X_t-1} Y(X_{t-i}+1,k_i)q(X_t-1,k_i) - \sum_{k_i=0}^{X_t} Y(X_{t-i},k_i)q(X_t,k_i) \right)}{P(X_t,X_{t-1},\ldots,X_{t-p})}
$$

$$
\geq - \frac{\frac{\phi_i X_{t-i}}{1+\alpha_i} \sum_{k_i=0}^{X_t} Y(X_{t-i},k_i)q(X_t,k_i)}{P(X_t,X_{t-1},\ldots,X_{t-p})}
$$

$$
\geq - \frac{X_{t-i}}{1+\alpha_i}
$$

and

$$
\frac{\partial \log P(X_t,X_{t-1},\ldots,X_{t-p})}{\partial \alpha_i} \leq \frac{\frac{\phi_i X_{t-i}}{1+\alpha_i} \sum_{k_i=0}^{X_t} \frac{k_i}{\alpha_i X_{t-i}} Y(X_{t-i},k_i)q(X_t,k_i)}{P(X_t,X_{t-1},\ldots,X_{t-p})}
$$

$$
\leq \frac{\frac{\phi_i X_t}{\alpha_i(1+\alpha_i)} \sum_{k_i=0}^{X_t} Y(X_{t-i},k_i)q(X_t,k_i)}{P(X_t,X_{t-1},\ldots,X_{t-p})} \leq \frac{X_t}{\alpha_i(1+\alpha_i)}, \quad (i = 2,\ldots,p).
$$

We aim to apply the results of Theorem 1.3 of Billingsley (1961) on estimates for the parameters. For this purpose, we have to impose some regularity conditions as follows:

(C1) The set $\{k : P(\varepsilon_t = m) = f(m,\lambda) = \frac{\lambda^m}{m!} e^{-\lambda} > 0\}$ does not depend on λ;
(C2) $E(\varepsilon_t^3) = \lambda^3 + 3\lambda^2 + \lambda < \infty$;
(C3) $P(\varepsilon_t = m)$ is three times continuously differentiable with respect to λ;
(C4) For any $\lambda' \in \mathcal{B}$, where $\mathcal{B}$ is an open subset of $\mathcal{R}$, there exists a neighbourhood $\mathcal{U}$ of λ' such that:
  (a) (1)$\sum_{k=0}^{\infty} \sup_{\lambda \in \mathcal{U}} f(k,\lambda) < \infty$,
  (b) (2)$\sum_{k=0}^{\infty} \sup_{\lambda \in \mathcal{U}} \left| \frac{\partial f(k,\lambda)}{\partial \lambda} \right| < \infty$,
  (c) (3)$\sum_{k=0}^{\infty} \sup_{\lambda \in \mathcal{U}} \left| \frac{\partial^2 f(k,\lambda)}{\partial \lambda^2} \right| < \infty$;
(C5) For any $\lambda' \in \mathcal{B}$ there exists a neighbourhood $\mathcal{U}$ of λ' and the sequence $\varphi_1(n) = \text{const } 1 \cdot n$, $\varphi_{11}(n) = \text{const } 2 \cdot n^2$, $\varphi_{111}(n) = \text{const } 3 \cdot n^3$, with suitable const 1, const 2, const 3 and $n \geq 0$ such that $\forall\, \lambda \in \mathcal{U}$ and $\forall\, m \leq n$, with $P(\varepsilon_t)$ nonvanishing $f(m,\lambda)$,

$$
\left| \frac{\partial f(k,\lambda)}{\partial \lambda} \right| \leq \varphi_1(n) f(m,\lambda), \quad \left| \frac{\partial^2 f(k,\lambda)}{\partial \lambda^2} \right| \leq \varphi_{11}(n) f(m,\lambda),
$$

$$\left|\frac{\partial^3 f(k,\lambda)}{\partial \lambda^3}\right| \le \varphi_{111}(n) f(m,\lambda),$$

and with respect to the stationary distribution of the process $\{X_t\}$,

$$E[\varphi_1^3(X_{p+1})] < \infty, \quad E[\varphi_{11}(X_{p+1})] < \infty,$$
$$E[\varphi_1(X_{p+1})\varphi_{11}(X_{p+1})] < \infty, \quad E[\varphi_{111}(X_{p+1})] < \infty,$$

(C6) Let $I(\boldsymbol{\theta}) = (\sigma_{ij})_{(2p+1)\times(2p+1)}$ denote the Fisher information matrix, i.e.

$$I(\boldsymbol{\theta}) = E\left(\frac{\partial \log P(X_{p+1},\ldots,X_1)}{\partial \theta_i} \cdot \frac{\partial \log P(X_{p+1},\ldots,X_1)}{\partial \theta_j}\right), \quad i,j = 1,\ldots,2p+1$$

and $I(\boldsymbol{\theta})$ is nonsingular.

Theorem 3.3: *Under Assumption* 2.1 *and the regularity conditions* (C1)–(C6)*, the CML estimate* $\hat{\boldsymbol{\theta}}_{CML}$ *is asymptotically normal, i.e.,*

$$\sqrt{N}(\hat{\boldsymbol{\theta}}_{CML} - \boldsymbol{\theta}) \xrightarrow{d} N(\mathbf{0}, I^{-1}(\boldsymbol{\theta})), \quad N \to \infty.$$

Proof: Theorem 3.3 is a special case of Theorem 1.3 of Billingsley (1961). Based on Assumption 2.1, Proposition 2.3 and the approach in Franke and Seligmann (1993), the theorem can be proved. And it is also discussed in Monteiro et al. (2012) that the above conditions (C1)-(C4) are all satisfied. (C5) follows by the definition of the Po-MDDRCINAR(p) model and the properties of Poisson distribution. Consequently, for the Po-MDDRCINAR(p) model it is just necessary to verify the last condition (analogous to condition (C6) in Franke and Seligmann (1993)) also holds. To this end, we need to verify that the following statements are all true.

(S_1) $E\left|\frac{\partial \log P(X_{p+1},\ldots,X_1)}{\partial \alpha_i}\right|^2 < \infty,$

(S_2) $E\left|\frac{\partial \log P(X_{p+1},\ldots,X_1)}{\partial \phi_i}\right|^2 < \infty,$

(S_3) $E\left|\frac{\partial \log P(X_{p+1},\ldots,X_1)}{\partial \lambda}\right|^2 < \infty,$

(S_4) $E\left|\frac{\partial \log P(X_{p+1},\ldots,X_1)}{\partial \alpha_i} \cdot \frac{\partial \log P(X_{p+1},\ldots,X_1)}{\partial \alpha_j}\right|^2 < \infty, i \neq j,$

(S_5) $E\left|\frac{\partial \log P(X_{p+1},\ldots,X_1)}{\partial \alpha_i} \frac{\partial \log P(X_{p+1},\ldots,X_1)}{\partial \phi_j}\right| < \infty,$

(S_6) $E\left|\frac{\partial \log P(X_{p+1},\ldots,X_1)}{\partial \alpha_i} \frac{\partial \log P(X_{p+1},\ldots,X_1)}{\partial \lambda}\right| < \infty,$

(S_7) $E\left|\frac{\partial \log P(X_{p+1},\ldots,X_1)}{\partial \phi_i} \frac{\partial \log P(X_{p+1},\ldots,X_1)}{\partial \phi_j}\right| < \infty, i \neq j,$

(S_8) $E\left|\frac{\partial \log P(X_{p+1},\ldots,X_1)}{\partial \phi_i} \frac{\partial \log P(X_{p+1},\ldots,X_1)}{\partial \lambda}\right| < \infty, \ i,j = 1,\ldots,p.$

We shall first prove Statement (S1). From Equations (12) and (13), we know that in the stationary state

$$E\left|\frac{\partial \log P(X_{p+1},\ldots,X_1)}{\partial \alpha_1}\right|^2 < C \cdot EX_p^2 < \infty, \tag{14}$$

and

$$E\left|\frac{\partial \log P(X_{p+1},\ldots,X_1)}{\partial \alpha_i}\right|^2 < C \cdot E(\max\{X_{p+1-i}, X_{p+1}\})^2 < \infty \tag{15}$$

for some suitable constant C.

Next, we will prove Statement (S2). As

$$\left|\frac{\partial \log P(X_{p+1},\ldots,X_1)}{\partial \phi_i}\right|$$

$$
= \begin{cases} \dfrac{\phi_1}{\phi_1 P(X_{p+1},\ldots,X_1)} \left[\displaystyle\sum_{k_1=0}^{\min(X_t,X_{t-1})} Z(X_{t-1},k_1)q(X_t,k_1) - q(X_t,k_0) \right], & j=1, \\ \dfrac{\phi_j}{\phi_j P(X_{p+1},\ldots,X_1)} \left[\displaystyle\sum_{k_j=0}^{X_t} Y(X_{t-j},k_j)q(X_t,k_j) - q(X_t,k_0) \right], & j=2,\ldots,p \end{cases}
$$

$$
< \frac{1}{\phi_i} < \infty, \quad i=1,\ldots,p,
$$

which implies

$$
E\left|\frac{\partial \log P(X_{p+1},\ldots,X_1)}{\partial \phi_i}\right|^2 < \infty. \tag{16}
$$

Then, we will prove Statement (S3). Based on (C5),

$$
\left|\frac{\partial \log P(X_{p+1},\ldots,X_1)}{\partial \lambda}\right| < \psi_1(X_{p+1}),
$$

which show that

$$
E\left|\frac{\partial \log P(X_{p+1},\ldots,X_1)}{\partial \lambda}\right|^2 < E\psi_1^2(X_{p+1}) < \infty. \tag{17}
$$

Finally, by Equations (14), (15), (16), (17) and (C5) it is concluded that Statements (S4)-(S8) hold. Therefore, the Fisher information matrix $I(\boldsymbol{\theta})$ is well-defined and according to (C6) it is nonsingular. ■

4. Simulation

In this section, we evaluate the small sample performance of our estimators via various simulations. Specifically, we anchor to check the finite-sample mean absolute deviation error (MADE) and mean squared errors (MSE) of alternative estimators as follows:

$$
\text{MADE} = \frac{1}{m}\sum_{i=1}^{m} |\hat{\theta}_i - \theta|, \quad \text{MSE} = \frac{1}{m}\sum_{i=1}^{m} (\hat{\theta}_i - \theta)^2,
$$

where $\hat{\theta}_i$ denotes any estimators of θ and m is the number of realizations. To this end, we generate simulation data from the Po-MDDRCINAR(2) model. According to the aforementioned definition of the model, as $\phi_0 + \phi_1 + \phi_2 = 1, 0 < \phi_i < 1, (i = 0, 1, 2), \phi_1\alpha_1 + \phi_2\alpha_2 < 1, 0 < \alpha_i < 1,\ i = 1, 2$ and each case under the study is non-degenerate and stationary. Thus, we first report representative parameter combinations for $\alpha_1 = \alpha_2 = 0.5,\ \lambda = 1$ based on the sample sizes $n = 100$ and 500 with 500 replications and the parameters are estimated using CLS and CML methods. The sum of ϕ_1 and ϕ_2 is confined within the range of (0.0,1.0). And, for each of these two parameters, different values range from 0.15 to 0.75, on a grid of 0.2. All possible parameter combinations of ϕ_1 and ϕ_2 are examined. Therefore, the results for MADE and MSE on different parameter combinations with $n = 100$ are displayed in Table 1, with similar results given with $n = 500$ in Table 2.

Moreover, values outside the allowed range for parameters might easily be obtained for small sample sizes. By Equation (11), the estimates of α_i and ϕ_i are decided by $\boldsymbol{\beta}$ and $\boldsymbol{\sigma}$. In Tables 1 and 2, such estimates have been adjusted by taking account that the restrictions on α_i and ϕ_i implying

$$
\begin{cases} 0 < \beta_1 + \beta_2 < 1, \\ \sigma_{11}\sigma_{22} > \beta_1^2\beta_2^2, \\ 0 < \sigma_{ii} < 0.25, \\ \sigma_{ii} < \beta_i(1-\beta_i), \quad i = 1, 2. \end{cases} \tag{18}
$$

Thus, if $\hat{\beta}_1 + \hat{\beta}_2 > 1$, $\hat{\beta}_i$ can be replaced by $\hat{\beta}_i/(\hat{\beta}_1 + \hat{\beta}_2)$, and other constraints on parameters can be made with similar adjustments.

Observing Tables 1 and 2, we can see that, for all considered parameter combinations, the MADE and MSE for the estimates of the parameters decrease as the sample size increases, which is in accordance with our expectation. Simultaneously, compared with CLS estimators, the result in Table 1 is that CML dominates CLS in terms of MADE

Table 1. MADE and MSE results for the Po-MDDRCINAR(2) model ($\alpha_1 = \alpha_2 = 0.5,\ \lambda = 1$) with sample size $n = 100$.

	ϕ_2								
	CLS					CML			
ϕ_1	0.15	0.35	0.55	0.75	ϕ_1	0.15	0.35	0.55	0.75
				MADE(α_1)					
0.15	0.2852	0.2811	0.2685	0.2604	0.15	0.1846	0.2073	0.2172	0.2289
0.35	0.2729	0.2591	0.2249		0.35	0.1443	0.1506	0.1543	
0.55	0.2377	0.2019			0.55	0.0997	0.0988		
0.75	0.2215				0.75	0.0615			
				MADE(α_2)					
0.15	0.2709	0.2327	0.1989	0.1765	0.15	0.1788	0.1518	0.1127	0.0814
0.35	0.2975	0.2447	0.1925		0.35	0.1940	0.1612	0.1244	
0.55	0.2782	0.2240			0.55	0.2043	0.1590		
0.75	0.2736				0.75	0.2116			
				MADE(ϕ_1)					
0.15	0.2590	0.2260	0.2113	0.1816	0.15	0.0814	0.0722	0.0894	0.0802
0.35	0.2821	0.2631	0.2260		0.35	0.1063	0.1176	0.1089	
0.55	0.2676	0.2326			0.55	0.1055	0.1030		
0.75	0.2099				0.75	0.0782			
				MADE(ϕ_2)					
0.15	0.3059	0.2877	0.2548	0.1922	0.15	0.0873	0.1153	0.1230	0.1024
0.35	0.3229	0.2717	0.2449		0.35	0.0725	0.1228	0.1409	
0.55	0.2572	0.2411			0.55	0.0944	0.1195		
0.75	0.2239				0.75	0.0934			
				MADE(λ)					
0.15	0.1613	0.1592	0.1790	0.2130	0.15	0.0756	0.0681	0.0788	0.0776
0.35	0.1505	0.1765	0.1841		0.35	0.0654	0.0841	0.1054	
0.55	0.1765	0.1793			0.55	0.0778	0.0851		
0.75	0.1833				0.75	0.0786			
				MSE(α_1)					
0.15	0.1039	0.1011	0.0922	0.0874	0.15	0.0493	0.0559	0.0624	0.0647
0.35	0.0970	0.0876	0.0699		0.35	0.0330	0.0348	0.0356	
0.55	0.0794	0.0594			0.55	0.0159	0.0154		
0.75	0.0705				0.75	0.0061			
				MSE(α_2)					
0.15	0.0943	0.0747	0.0553	0.0437	0.15	0.0453	0.0342	0.0204	0.0102
0.35	0.1108	0.0829	0.0544		0.35	0.0523	0.0379	0.0245	
0.55	0.1002	0.0715			0.55	0.0553	0.0375		
0.75	0.0973				0.75	0.0586			
				MSE(ϕ_1)					
0.15	0.1324	0.1036	0.0883	0.0689	0.15	0.0123	0.0168	0.0290	0.0187
0.35	0.1122	0.0981	0.0749		0.35	0.0178	0.0247	0.0205	
0.55	0.0957	0.0771			0.55	0.0187	0.0176		
0.75	0.0682				0.75	0.0102			
				MSE(ϕ_2)					
0.15	0.1743	0.1206	0.0878	0.5584	0.15	0.0130	0.0196	0.0241	0.0168
0.35	0.1875	0.1095	0.0823		0.35	0.0104	0.0246	0.0314	
0.55	0.1254	0.0862			0.55	0.0310	0.0242		
0.75	0.1006				0.75	0.0301			
				MSE(λ)					
0.15	0.0421	0.0425	0.0554	0.0757	0.15	0.0092	0.0076	0.0097	0.0099
0.35	0.0378	0.0509	0.0560		0.35	0.0068	0.0105	0.0164	
0.55	0.0475	0.0519			0.55	0.0096	0.0114		
0.75	0.0541				0.75	0.0096			

and MSE for all considered parameter combinations. And, the similar results can be obtained for the MADE and MSE of the parameter estimation in Table 2. According to the aforementioned analysis, the parameter estimations for these two estimation methods perform better with sample size increasing, which indicates that a large sample size may be needed to obtain reasonable results.

For a more comprehensive comparison of the MSE of the CML and CLS for parameters, we again report four samples for parameter values of $\alpha_i, \phi_i,\ i = 1, 2$ and λ, namely,

(1) Sample 1: $\alpha_1 = 0.5,\ \alpha_2 = 0.5,\ \phi_1 = 0.25,\ \phi_2 = 0.25$ and $\lambda = 1$,
(2) Sample 2: $\alpha_1 = 0.5,\ \alpha_2 = 0.5,\ \phi_1 = 0.25,\ \phi_2 = 0.25$ and $\lambda = 2$,
(3) Sample 3: $\alpha_1 = 0.6,\ \alpha_2 = 0.7,\ \phi_1 = 0.4,\ \phi_2 = 0.5$ and $\lambda = 1$,
(4) Sample 4: $\alpha_1 = 0.6,\ \alpha_2 = 0.7,\ \phi_1 = 0.4,\ \phi_2 = 0.5$ and $\lambda = 2$.

Figure 1 gives the paths for these four typical samples in the Po-MDDRCINAR(2) model with a sample size 100. The representative results for Samples 1, 2, 3 and 4 are summarized in Table 3 based on the CLS and CML estimation

Table 2. MADE and MSE results for the Po-MDDRCINAR(2) Model ($\alpha_1 = \alpha_2 = 0.5$, $\lambda = 1$) with sample size $n = 500$.

	ϕ_2								
	CLS					CML			
ϕ_1	0.15	0.35	0.55	0.75	ϕ_1	0.15	0.35	0.55	0.75
				MADE(α_1)					
0.15	0.2686	0.2415	0.2228	0.1982	0.15	0.1631	0.1535	0.1400	0.1366
0.35	0.2109	0.1521	0.1256		0.35	0.0774	0.0752	0.0716	
0.55	0.1507	0.1091			0.55	0.0457	0.0440		
0.75	0.1229				0.75	0.0334			
				MADE(α_2)					
0.15	0.2453	0.1514	0.1004	0.0903	0.15	0.1276	0.0731	0.0492	0.0405
0.35	0.2549	0.1322	0.0884		0.35	0.1459	0.0777	0.0512	
0.55	0.2351	0.1267			0.55	0.1314	0.0724		
0.75	0.2170				0.75	0.1359			
				MADE(ϕ_1)					
0.15	0.1802	0.1517	0.1508	0.1162	0.15	0.0504	0.0562	0.0580	0.0498
0.35	0.1836	0.1452	0.1096		0.35	0.0560	0.0628	0.0573	
0.55	0.1565	0.1191			0.55	0.0526	0.0504		
0.75	0.1329				0.75	0.0452			
				MADE(ϕ_2)					
0.15	0.2013	0.1512	0.1188	0.1058	0.15	0.0482	0.0622	0.0617	0.0572
0.35	0.1771	0.1274	0.1018		0.35	0.0615	0.0644	0.0610	
0.55	0.1637	0.1255			0.55	0.0550	0.0605		
0.75	0.1500				0.75	0.0561			
				MADE(λ)					
0.15	0.0714	0.0738	0.0820	0.0985	0.15	0.0364	0.0420	0.0450	0.0475
0.35	0.0728	0.0844	0.0902		0.35	0.0403	0.0472	0.0528	
0.55	0.0724	0.0906			0.55	0.0434	0.0469		
0.75	0.0768				0.75	0.0447			
				MSE(α_1)					
0.15	0.0945	0.0770	0.0696	0.0569	0.15	0.0405	0.0365	0.0301	0.0293
0.35	0.0639	0.0361	0.0259		0.35	0.0095	0.0086	0.0079	
0.55	0.0347	0.0188			0.55	0.0033	0.0305		
0.75	0.0263				0.75	0.0017			
				MSE(α_2)					
0.15	0.0813	0.0335	0.0152	0.0122	0.15	0.0261	0.0086	0.0039	0.0026
0.35	0.0873	0.0271	0.0123		0.35	0.0328	0.0095	0.0042	
0.55	0.0756	0.0247			0.55	0.0287	0.0083		
0.75	0.0658				0.75	0.0300			
				MSE(ϕ_1)					
0.15	0.0731	0.0505	0.0511	0.0265	0.15	0.0039	0.0109	0.0127	0.0081
0.35	0.0620	0.0384	0.0204		0.35	0.0057	0.0087	0.0072	
0.55	0.0396	0.0223			0.55	0.0050	0.0044		
0.75	0.0307				0.75	0.0031			
				MSE(ϕ_2)					
0.15	0.0920	0.0405	0.0224	0.0172	0.15	0.0035	0.0060	0.0065	0.0051
0.35	0.0716	0.0290	0.0167		0.35	0.0099	0.0070	0.0063	
0.55	0.0618	0.0253			0.55	0.0077	0.0070		
0.75	0.0479				0.75	0.0106			
				MSE(λ)					
0.15	0.0080	0.0087	0.0108	0.0153	0.15	0.0020	0.0027	0.0029	0.0034
0.35	0.0084	0.0113	0.0127		0.35	0.0024	0.0034	0.0043	
0.55	0.0085	0.0132			0.55	0.0028	0.0035		
0.75	0.0096				0.75	0.0030			

methods, where the negative estimates have also been adjusted according to Equation (18). From the simulation results, we can see that, as the sample size n increases, all the estimates seem reasonable with CML being more efficient.

Meanwhile, Figure 2 displays the Q-Q plots for these four typical samples when $n = 1000$. The Q-Q plot visually shows that the CLS and CML estimators are asymptotically normal for all the parameters, and that the distribution of estimators obtained by the CML method is apparently closer to the normal distribution. Thereby, the aforementioned analysis is analogous with the theorem of asymptotic normality of parameter estimators.

5. COVID-19 data on suspected cases in China

In this section, in order to illustrate the application of the proposed Po-MDDRCINAR(p) model based on binomial and negative binomial thinning operators with Poisson innovation, we analyse the COVID-19 data on suspected cases in China, which can be obtained from the website of National Health Commission of the People's Republic of China (*http://www.nhc.gov.cn/xcs/yqtb/listgzbd.shtml*). Moreover, the performance of the model is evaluated in

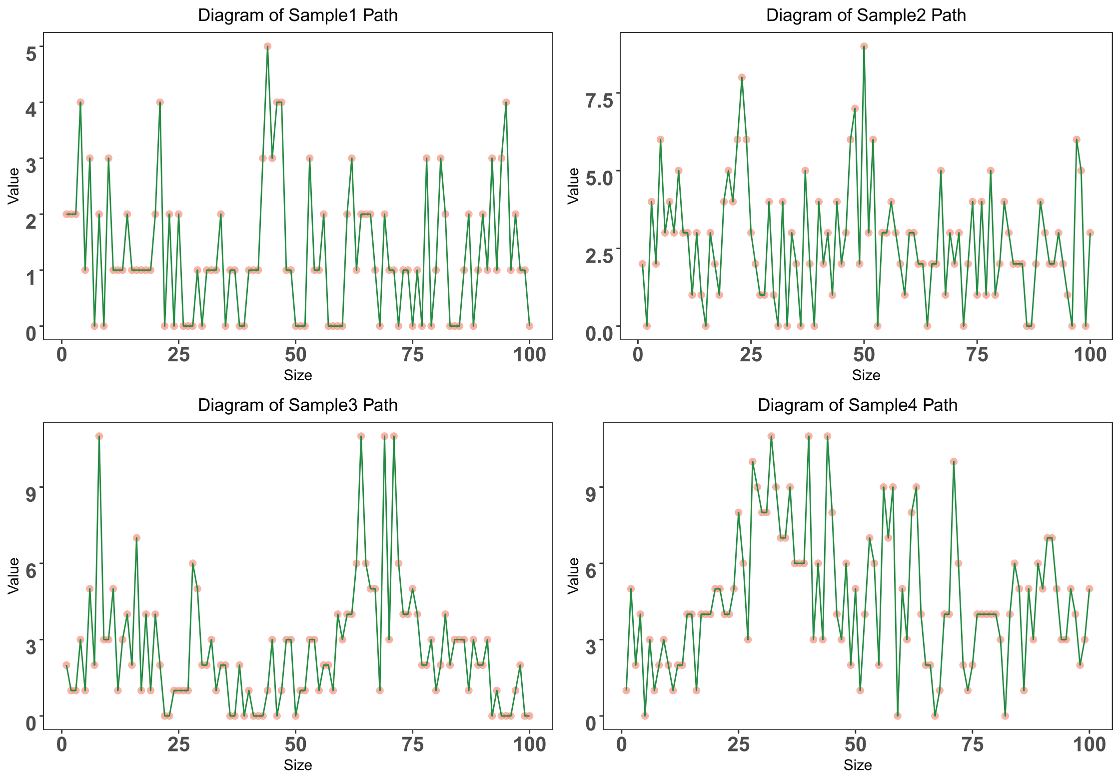

Figure 1. Paths for Samples 1, 2, 3 and 4 of the Po-MDDRCINAR(2) model.

terms of four criteria: Akaike information criterion (AIC), Bayesian information criterion (BIC), Hannan–Quinn criterion (HQ) and the root mean squares of differences of observations and predicted values (RMS).

The analysed count set is about the COVID-19 data on suspected cases in China, which consist of 191 daily observations in a period from August 29, 2021 to March 7, 2022. Note that the mean of the data is approximately 3.0785, and the variance is around 8.1149. Moreover, the plots of the time series, its empirical autocorrelation (ACF) and partial autocorrelation function (PACF) are shown in Figures 3, 4 and 5, respectively. Inspecting the plot of the PACF in more detail, however, an autoregressive model including X_{t-1}, X_{t-2} and ε_t appears to be reasonable. Hence, we consider three candidate models to analyse the data, which are as follows.

Model I:

$$X_t = \alpha_{t1} \circ X_{t-1} + \alpha_{t,2} * X_{t-2} + \varepsilon_t \tag{19}$$

with the joint distribution of $\{\alpha_{t1}, \alpha_{t2}\}$ given by

$$\begin{cases} P(\alpha_{t1} = \alpha_1, \alpha_{t2} = 0) = \phi_1; \\ P(\alpha_{t1} = 0, \alpha_{t2} = \alpha_2) = \phi_2; \\ P(\alpha_{t1} = 0, \alpha_{t2} = 0) = 1 - \phi_1 - \phi_2. \end{cases} \tag{20}$$

And, $\{\varepsilon_t\}$ is an i.i.d. Poisson sequence with mean λ and $\alpha_1,\ \alpha_2 \in [0, 1)$.

Model II:

$$X_t = \alpha_{t1} \circ X_{t-1} + \alpha_{t2} \circ X_{t-2} + \varepsilon_t, \tag{21}$$

where the joint distribution of $\{\alpha_{t1}, \alpha_{t2}\}$ is the same as Equation (20) and $\{\varepsilon_t\}$ is an i.i.d. Poisson sequence with mean λ.

Model III:

$$X_t = \begin{cases} \alpha \circ_t X_{t-1} + \varepsilon_t, & w.p.\ \phi_1, \\ \alpha *_t X_{t-2} + \varepsilon_t, & w.p.\ 1 - \phi_1, \end{cases} \tag{22}$$

where $\{\varepsilon_t\}$ is an i.i.d. Poisson sequence with mean λ and $\alpha \in (0, 1)$.

Table 4 displays the estimates of the parameters based on the CML estimation method provided in Section 3. We also present the associated AIC, BIC, HQ and RMS in this table. Certainly, we should consider further checks

Table 3. Mean values and MSE (within parentheses) of estimates for parameters in Samples 1, 2, 3 and 4.

n	α_1	α_2	ϕ_1	ϕ_2	λ
			$\alpha_1=\alpha_2=0.5, \phi_1=\phi_2=0.25, \lambda=1$		
			CLS		
100	0.4662(0.0979)	0.4101(0.0910)	0.2804(0.0986)	0.4033(0.1441)	1.0640(0.0451)
300	0.4493(0.0823)	0.4148(0.0735)	0.2931(0.0796)	0.3499(0.0872)	1.0205(0.0143)
500	0.4183(0.0662)	0.4277(0.0553)	0.3122(0.0604)	0.3151(0.0598)	1.0130(0.0079)
800	0.4259(0.0485)	0.4511(0.0343)	0.3194(0.0562)	0.2780(0.0306)	1.0050(0.0049)
1000	0.4386(0.0482)	0.4504(0.0320)	0.3108(0.0516)	0.2662(0.0235)	1.0144(0.0044)
			CML		
100	0.3769(0.0486)	0.4398(0.0412)	0.1875(0.0158)	0.2202(0.0148)	0.9822(0.0079)
300	0.3793(0.0304)	0.4061(0.0278)	0.1803(0.0095)	0.1831(0.0098)	0.9590(0.0042)
500	0.4082(0.0178)	0.4153(0.0197)	0.1970(0.0067)	0.1913(0.0075)	0.9648(0.0029)
800	0.4265(0.0115)	0.4375(0.0102)	0.2086(0.0057)	0.2027(0.0047)	0.9701(0.0019)
1000	0.4342(0.0087)	0.4442(0.0084)	0.2104(0.0036)	0.2066(0.0039)	0.9751(0.0014)
			$\alpha_1=\alpha_2=0.5, \phi_1=\phi_2=0.25, \lambda=2$		
			CLS		
100	0.4342(0.0866)	0.4432(0.0825)	0.2984(0.0889)	0.3533(0.1137)	2.1118(0.1611)
300	0.4856(0.0615)	0.4381(0.0592)	0.2736(0.0547)	0.3146(0.0698)	2.0291(0.0457)
500	0.4432(0.0440)	0.4849(0.0393)	0.2798(0.0394)	0.2582(0.0337)	2.0318(0.0334)
800	0.4749(0.0282)	0.4807(0.0237)	0.2497(0.0197)	0.2532(0.0181)	2.0177(0.0171)
1000	0.4596(0.0219)	0.4818(0.0180)	0.2672(0.0192)	0.2481(0.0122)	2.0060(0.0157)
			CML		
100	0.3798(0.0408)	0.3895(0.0396)	0.1775(0.0146)	0.1775(0.0146)	1.9163(0.0182)
300	0.4111(0.0192)	0.4091(0.0205)	0.1957(0.0088)	0.1865(0.0081)	1.9215(0.0104)
500	0.4314(0.0109)	0.4400(0.0107)	0.2074(0.0057)	0.2048(0.0072)	1.9422(0.0063)
800	0.4465(0.0069)	0.4536(0.0060)	0.2112(0.0023)	0.2134(0.0051)	1.9512(0.0042)
1000	0.4469(0.0058)	0.4598(0.0047)	0.2172(0.0021)	0.2147(0.0028)	1.9532(0.0037)
			$\alpha_1=0.6, \alpha_2=0.7, \phi_1=0.4, \phi_2=0.5, \lambda=1$		
			CLS		
100	0.4496(0.0706)	0.5182(0.0683)	0.4177(0.0628)	0.4760(0.0545)	1.1380(0.1034)
300	0.4688(0.0297)	0.5625(0.0275)	0.3968(0.0182)	0.5042(0.0158)	1.0488(0.0315)
500	0.4676(0.0255)	0.5785(0.0206)	0.3935(0.0123)	0.5016(0.0104)	1.0391(0.0217)
800	0.4667(0.0229)	0.5878(0.0167)	0.3965(0.0082)	0.4961(0.0066)	1.0254(0.0130)
1000	0.4664(0.0217)	0.5888(0.0157)	0.4017(0.0057)	0.4966(0.0049)	1.0192(0.0093)
			CML		
100	0.5078(0.0276)	0.6057(0.0321)	0.3214(0.0136)	0.4072(0.0174)	0.9261(0.0197)
300	0.5592(0.0066)	0.6592(0.0078)	0.3577(0.0051)	0.4510(0.0038)	0.9513(0.0071)
500	0.5746(0.0033)	0.6710(0.0039)	0.3734(0.0045)	0.4666(0.0027)	0.9679(0.0038)
800	0.5786(0.0020)	0.6772(0.0027)	0.3761(0.0016)	0.4725(0.0018)	0.9711(0.0029)
1000	0.5826(0.0015)	0.6785(0.0022)	0.3780(0.0011)	0.4744(0.0011)	0.9761(0.0021)
			$\alpha_1=0.6, \alpha_2=0.7, \phi_1=0.4, \phi_2=0.5, \lambda=2$		
			CLS		
100	0.4632(0.0477)	0.5562(0.0438)	0.4123(0.0483)	0.4706(0.0472)	2.2394(0.3601)
300	0.4714(0.0228)	0.5799(0.0198)	0.3937(0.0140)	0.4863(0.0131)	2.1025(0.1292)
500	0.4683(0.0216)	0.5888(0.0161)	0.3922(0.0101)	0.4916(0.0088)	2.0722(0.0751)
800	0.4697(0.0192)	0.5899(0.0139)	0.3919(0.0053)	0.4946(0.0050)	2.0443(0.0411)
1000	0.4677(0.0190)	0.5930(0.0132)	0.5969(0.0043)	0.4926(0.0040)	2.0293(0.0384)
			CML		
100	0.5575(0.0125)	0.6487(0.0163)	0.3512(0.0077)	0.4373(0.0076)	1.8883(0.0447)
300	0.5799(0.0034)	0.6770(0.0033)	0.3710(0.0028)	0.4663(0.0018)	1.9335(0.0162)
500	0.5821(0.0020)	0.6819(0.0022)	0.3781(0.0018)	0.4756(0.0016)	1.9457(0.0102)
800	0.5870(0.0010)	0.6862(0.0013)	0.3831(0.0017)	0.4798(0.0009)	1.9531(0.0064)
1000	0.5923(0.0008)	0.6869(0.0010)	0.3851(0.0006)	0.4850(0.0005)	1.9630(0.0049)

on the appropriateness of the model before making any decision as Jung et al. (2006), e.g. regarding the stochastic properties of the fitted model, or with to properties of the standardized Pearson residuals:

$$e_t = \frac{X_t - E(X_t \mid X_{t-1}, X_{t-2})}{\sqrt{\mathrm{Var}(X_t \mid X_{t-1}, X_{t-2})}} \quad \text{for } t = 3, \ldots, T. \tag{23}$$

If the model is correctly specified, then these residuals should have zero mean and unit variance, namely, the distribution of the residual is apparently closer to the standard normal distribution. Thus, the plots and Q-Q plots of standardized Pearson residuals based the Models I, II and III are provided in Figure 6.

Based on the AIC, BIC, HQ and RMS results, one may choose the Model I for this data. On the other hand, because the mean and variance of the standardized Pearson residuals of Models I, II and III are (0.0080,1.1009), (−0.0180,1.2522) and (0.1275,1.2035), respectively, and the Q-Q plots in Figure 6, which show that the distribution of the standardized Pearson residual of Model I is obviously closer to the standard normal distribution than the other two models. Thus, one may again choose Model I for this data based on the above analysis.

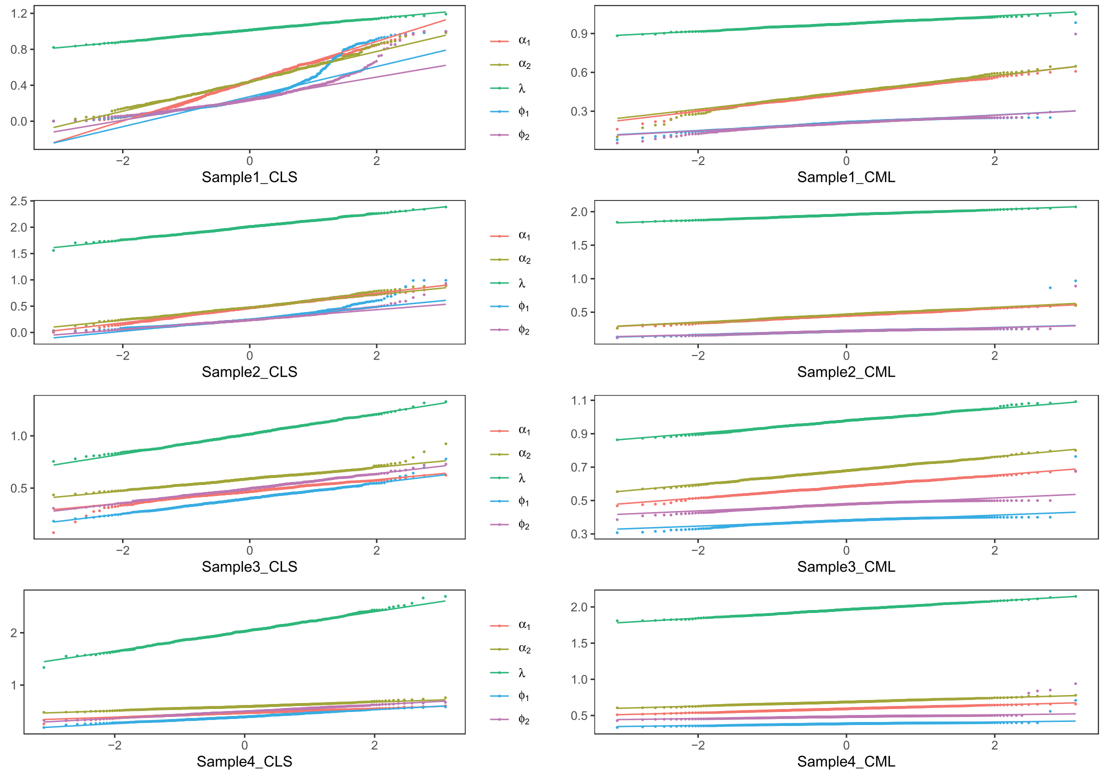

Figure 2. The Q-Q plots of the estimates based on CLS on the left side and CML on the right side.

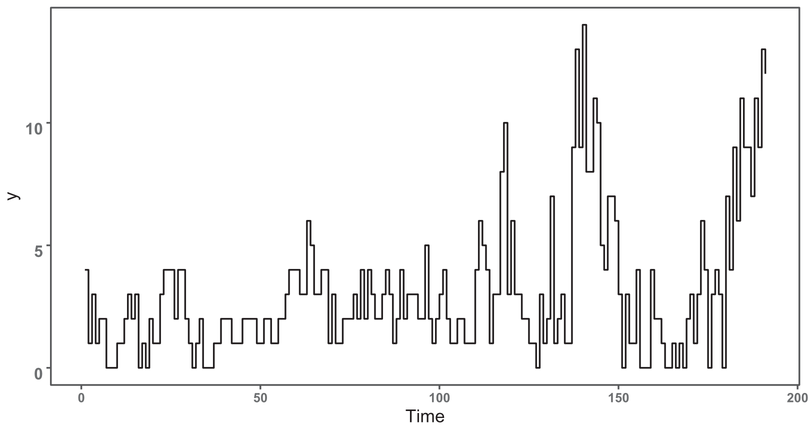

Figure 3. The path diagram of the COVID-19 data on suspected cases in China.

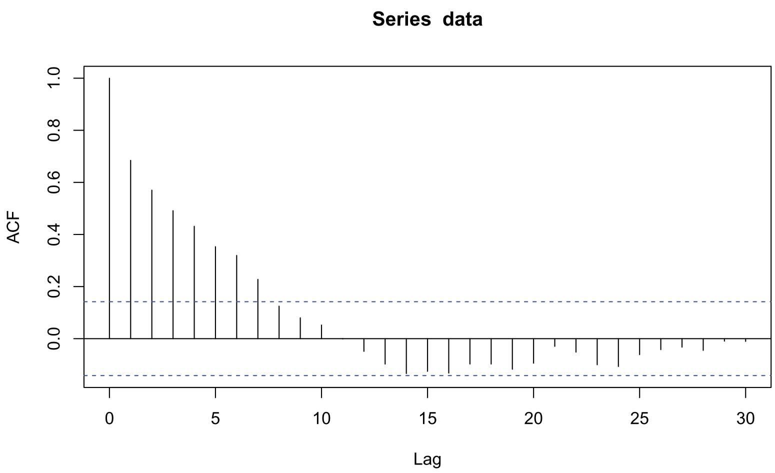

Figure 4. ACF plot of the COVID-19 data on suspected cases in China.

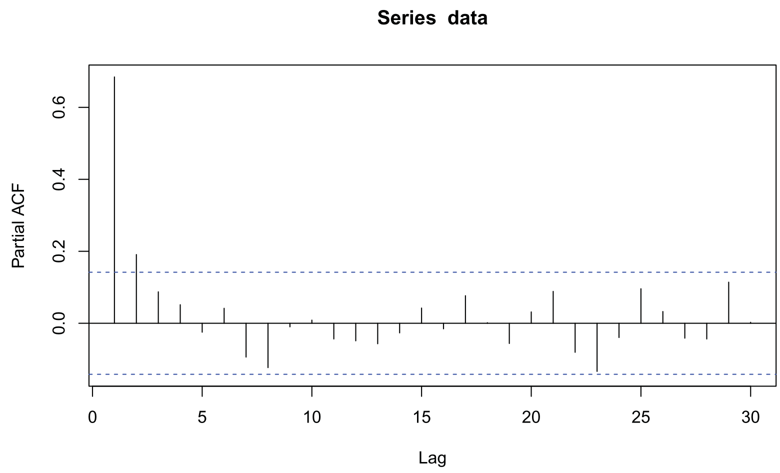

Figure 5. PACF plot of the COVID-19 data on suspected cases in China.

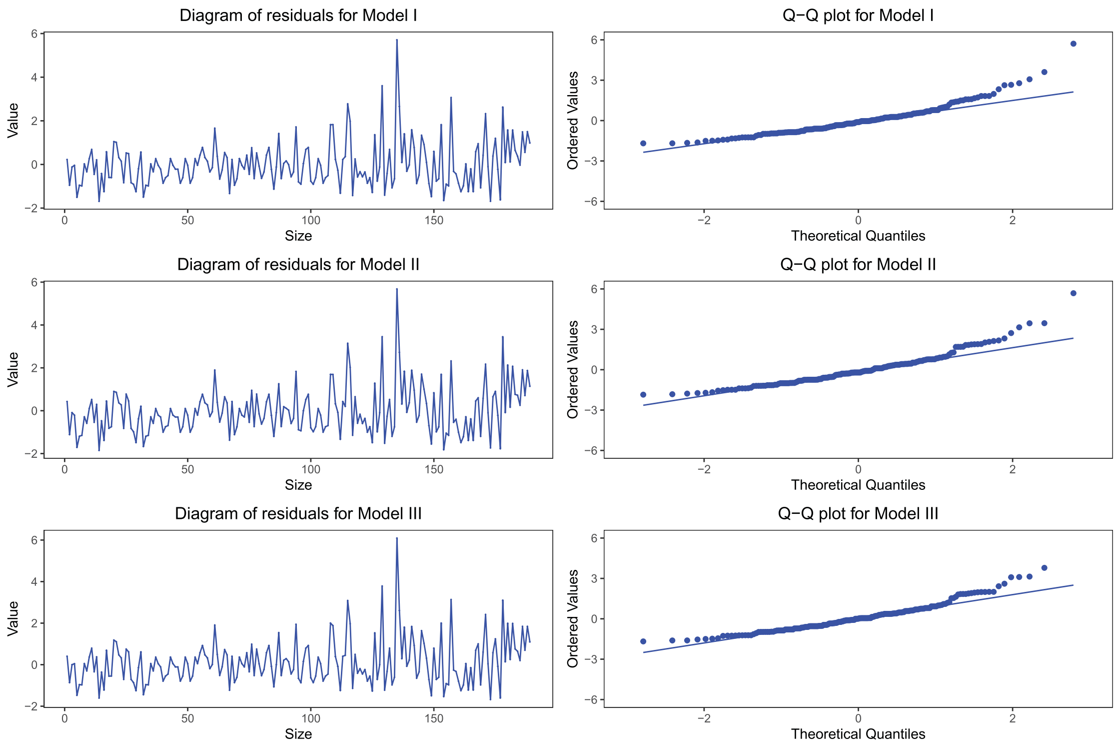

Figure 6. Q-Q plots of the standardized Pearson residuals on models I-III.

Table 4. Maximum-likelihood estimation of parameters, AIC, BIC, HQ and RMS for COVID-19 data on suspected cases in China.

	Parameters						Criteria			
Model	α	α_1	α_2	ϕ_1	ϕ_2	λ	AIC	BIC	HQ	RMS
I		0.6650	0.8228	0.6004	0.3426	0.9728	729.63	732.88	729.28	1.999
II		0.6018	0.9921	0.6866	0.1320	1.3263	738.61	741.87	738.27	2.073
III	0.6854			0.5574		0.9471	732.41	735.66	732.07	2.058

6. Summary and conclusion

In this paper, we have introduced a pth-order mixed dependence-driven random coefficient integer-valued autoregressive model (Po-MDDRCINAR(p)) with Poisson innovation based on binomial thinning and negative binomial thinning operators. The practical relevance of the Po-MDDRCINAR(p) model was demonstrated using a COVID-19 data set of suspected cases in China. Important statistical properties of the Po-MDDRCINAR(p) model were derived, the strict stationarity and ergodicity properties of the process are established, CLS and CML approaches for parameter estimation were discussed and their asymptotic properties are derived.

The simulation results confirm the effectiveness of our methods. Finally, a real data analysis demonstrates that the Po-MDDRCINAR(p) model performs best on the COVID-19 data set. Our research focuses on the establishment of the nonlinear time series model based on binomial and negative binomial thinning operators, as well as the parameter estimation problems. In future studies, we plan to investigate the forecasting problem of the Po-MDDRCINAR(p) model. Moreover, we also plan to take into consideration of the detailed inference for a pth-order serially dependent mixed dependence-driven random coefficient integer-valued autoregressive model (SD-MDDRCINAR(p)), where the mean of the innovations is linearly increased by the population size at time $\{t-1, t-2, \ldots, t-p\}$, to deal with the overdispersed data.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work is supported by Fundamental Research Program of Shanxi Province [Grant No. 202103021223084], National Natural Science Foundation of China [Grant No. 12271231], Natural Science Foundation of Henan Province [Grant No. 222300420127] and 2023 Graduate Education Innovation Program Various Projects and Special Funds for Innovation and Entrepreneurship [Grant No. RC2300003332].

ORCID

Xiufang Liu http://orcid.org/0000-0002-1385-1619

References

Al-Osh, M. A., & Alzaid, A. A. (1987). First-order integer-valued autoregressive (INAR(1)) process. *Journal of Time Series Analysis*, *8*(3), 261–275. https://doi.org/10.1111/j.1467-9892.1987.tb00438.x

Al-Osh, M. A., & Aly, E. E. A. A. (1992). First order autoregressive time series with negative binomial and geometric marginals. *Communications in Statistics–Theory and Methods*, *21*(9), 2483–2492. https://doi.org/10.1080/03610929208830925

Alzaid, A. A., & Al-Osh, M. A. (1988). First-order integer-valued autoregressive (INAR(1)) processes: Distributional and regression properties. *Statistica Neerlandica*, *42*(1), 53–61. https://doi.org/10.1111/j.1467-9574.1988.tb01521.x

Alzaid, A. A., & Al-Osh, M. A. (1990). An integer-valued *p*th-order autoregressive structure (INAR(p)) process. *Journal of Applied Probability*, *27*(2), 314–324. https://doi.org/10.2307/3214650

Billingsley, P. (1961). *Statistical Inference for Markov Processes*. University of Chicago Press.

Brockwell, P. J., & Davis, R. A. (1987). *Time Series: Theory and Methods*. Springer.

Du, J. G., & Li, Y. (1991). The integer-valued autoregressive (INAR(p)) model. *Journal of Time Series Analysis*, *12*(2), 129–142. https://doi.org/10.1111/j.1467-9892.1991.tb00073.x

Franke, J., & Seligmann, T. (1993). Conditional maximum likelihood estimates for INAR(1) processes and their application to modeling epileptic seizure counts. In T. Subba Rao (Ed.), *Developments in Time Series Analysis* (pp. 310–330). Chapman Hall. https://doi.org/10.1007/978-1-4899-4515-0_22

Hwang, S. Y., & Basawa, I. V. (1998). Parameter estimation for generalized random coefficient autoregressive processes. *Journal of Statistical Planning and Inference*, *68*(2), 323–337. https://doi.org/10.1016/S0378-3758(97)00147-X

Jung, R. C., Kukuk, M., & Liesenfeld, R. (2006). Time series of count data: Modeling, estimation and diagnostics. *Computational Statistics and Data Analysis*, *51*(4), 2350–2364. https://doi.org/10.1016/j.csda.2006.08.001

Klimko, L. A., & Nelson, P. I. (1978). On conditional least squares estimation for stochastic processes. *The Annals of Statistics*, *6*(3), 629–642. https://doi.org/10.1214/aos/1176344207

Liu, X. F., Jiang, H., & Wang, D. H. (2023). Estimation of parameters in the MDDRCINAR(*p*) model. *Journal of Statistical Computation and Simulation*, *93*(6), 983–1010. https://doi.org/10.1080/00949655.2021.1970163

Liu, X. F., & Wang, D. H. (2019). Estimation of parameters in the DDRCINAR(*p*) model. *Brazilian Journal of Probability and Statistics*, *33*(3), 638–673. https://doi.org/10.1214/18-BJPS405

Liu, X. F., Wang, D. H., Deng, D. L., Cheng, J. H., & Lu, F. L. (2020). Maximum likelihood estimation of the DDRCINAR(*p*) model. *Communications in Statistics-Theory and Methods*, *50*(24), 6231–6255. https://doi.org/10.1080/03610926.2020.1741627

Monteiro, M., Scotto, M. G., & Pereira, I. (2012). Integer-valued self-exciting threshold autoregressive processes. *Communications in Statistics-Theory and Methods*, *41*(15), 2717–2737. https://doi.org/10.1080/03610926.2011.556292

Nastić, A. S., Ristić, M. M., & Bakouch, H. S. (2012). A combined geometric INAR(*p*) model based on negative binomial thinning. *Mathematical and Computer Modelling*, *55*(5-6), 1665–1672. https://doi.org/10.1016/j.mcm.2011.10.080

Nicholls, D. F., & Quinn, B. G. (1982). *Random Coefficient Autoregressive Models: An Introduction*. Springer Science and Business Media.

Ristić, M. M., Bakouch, H. S., & Nastić, A. S. (2009). A new geometric first-order integer-valued autoregressive (NGINAR(1)) process. *Statistical Planning and Inference*, *139*(7), 2218–2226. https://doi.org/10.1016/j.jspi.2008.10.007

Ristić, M. M., & Nastić, A. S. (2012). A mixed INAR(*p*) model. *Journal of Time Series Analysis*, *33*(6), 903–915. https://doi.org/10.1111/j.1467-9892.2012.00806.x

Shirozhan, M., & Mohammadpour, M. (2020). An INAR(1) model based on the Pegram and thinning operators with serially dependent innovation. *Communications in Statistics–Simulation and Computation*, *49*(10), 2617–2638. https://doi.org/10.1080/03610918.2018.1521975

Steutel, F. W., & Van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *Annals of Probability*, *7*(5), 893–899. https://doi.org/10.1214/aop/1176994950

Weiß, C. H. (2008). The combined INAR(p) models for time series of counts. *Statistics and Probability Letters*, *78*(13), 1817–1822. https://doi.org/10.1016/j.spl.2008.01.036

Yu, M. J., Wang, D. H., & Yang, K. (2019). A class of observation-driven random coefficient INAR(1) processes based on negative binomial thinning. *Journal of the Korean Statistical Society*, *48*(2), 248–264. https://doi.org/10.1016/j.jkss.2018.11.004

Zhang, H. X., & Wang, D. H. (2015). Inference for random coefficient INAR(1) process based on frequency domain analysis. *Communications in Statistics-Simulation and Computation*, *44*(4), 1078–1100. https://doi.org/10.1080/03610918.2013.804556

Zheng, H. T., & Basawa, I. V. (2008). First-order observation-driven integer-valued autoregressive processes. *Statistics & Probability Letters*, *78*(1), 1–9. https://doi.org/10.1016/j.spl.2007.04.017

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

Zheng, H. T., Ishwar, V., Basawa, D., & Somnath, D. (2007). First-order random coefficient integer-valued autoregressive processes. *Journal of Statistical Planning and Inference*, *137*(1), 212–229. https://doi.org/10.1016/j.jspi.2005.12.003