Brazilian Journal of Probability and Statistics
2019, Vol. 33, No. 3, 638–673
https://doi.org/10.1214/18-BJPS405
© Brazilian Statistical Association, 2019

Estimation of parameters in the DDRCINAR(p) model

Xiufang Liu and Dehui Wang
Jilin University

Abstract. This paper discusses a pth-order dependence-driven random coefficient integer-valued autoregressive time series model (DDRCINAR(p)). Stationarity and ergodicity properties are proved. Conditional least squares, weighted least squares and maximum quasi-likelihood are used to estimate the model parameters. Asymptotic properties of the estimators are presented. The performances of these estimators are investigated and compared via simulations. In certain regions of the parameter space, simulative analysis shows that maximum quasi-likelihood estimators perform better than the estimators of conditional least squares and weighted least squares in terms of the proportion of within-Ω estimates. At last, the model is applied to two real data sets.

1 Introduction

Integer-valued time series models have received growing attention recently. These models can be broadly classified into two types. One is regression-type models, and the other is 'thinning' models. (See Davis, Dunsmuir and Wang (1999), for a review of the regression models.) Steutel and Van Harn (1979) gave the 'thinning' operator '$\circ$', which has been used by many authors. For example, Al-Osh and Alzaid (1987, 1988, 1990) have studied the thinning models, as well as Du and Li (1991), Latour (1997, 1998), Brannas and Hellstrom (2001) and Li, Wang and Zhang (2015), in a slightly more general form, among others. The work presented in this paper is that we propose a dependence-driven random coefficient thinning model for the pth-order integer-valued autoregression.

The first-order integer-valued autoregressive (INAR(1)) process is introduced by Al-Osh and Alzaid (1987). It is defined by

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

where

$$\phi \circ X_{t-1} = \sum_{i=1}^{X_{t-1}} B_{1,t},$$

here, the so-called counting series $\{B_{1,t}\}$ are independent and identically distributed (i.i.d.) Bernoulli random variables with success probability $\phi \in [0, 1]$ and

Key words and phrases. Conditional least squares, maximum quasi-likelihood, DDRCINAR(p) model, weighted conditional least squares, asymptotic distribution.

Received March 2018; accepted May 2018.

$\{\varepsilon_t\}$ is a sequence of i.i.d. non-negative integer-valued random variables and independent of the counting series $\{B_{1,t}\}$. Thus, $\phi \circ X_{t-1}$ is a binomial random variable with ϕ and X_{t-1} as parameters, namely $\phi \circ X_{t-1} \sim B(X_{t-1}, \phi)$.

In many real-life situations there is a INAR model to be used for some situations, which could be found in reliability theory, meteorology, insurance theory, communications, medicine, law and social sciences, such as counts of accidents, detected errors, transmitted messages, patients, crime victimization, etc. As an example of a standard INAR(1) model, let X_t denote the number of surviving epileptic patients in a hospital at time t, ϕ the probability of survival from time $t-1$ to t, and ε_t the number of new epileptic patients admitted at time t. As further example, suppose X_t denote the number of unemployed in the tth month. Then X_t can be modeled as the sum of the previously unemployed $\phi \circ X_{t-1}$ and the newly unemployed ε_t. In particular, the parameter ϕ may vary with time and it may be random as the survival rate (the unemployment rate) ϕ may be affected by various environmental factors, such as the quality of health care, the state of health of patients, etc. (affected by factors such as the state of the economy, productivity growth, etc.). Thus, it is necessary to research random-coefficient INAR model. Recently, Zheng, Basawa and Datta (2007) and Zheng and Basawa (2008) studied the random-coefficient INAR(1) model as well as Zhang and Wang (2015).

The pth-order integer-valued autoregressive (INAR(p)) model is recursively defined by Du and Li (1991) as

$$X_t = \sum_{i=1}^{p} \phi_i \circ X_{t-i} + \varepsilon_t, \qquad t \geq 1,$$

where, for $i = 1, \dots, p$,

$$\phi_i \circ X_{t-i} = \sum_{i=1}^{X_{t-i}} B_{i,t},$$

here $\{B_{i,t}\}$, $i \in \{1, \dots, p\}$ are independent Bernoulli-distributed variables, where $\{B_{i,t}\}$ has success probability $\phi_i \in [0, 1]$ and $\{\varepsilon_t\}$ is a sequence of i.i.d. non-negative integer-valued random variables and independent of all the counting series and $\sum_{i=1}^{p} \phi_i < 1$.

In the specification of Du and Li (1991) and Latour (1998), the autocorrelation structure of an INAR(p) process was the same as that of an AR(p) process. Du and Li (1991) proved the existence and ergodic property of the INAR(p) model, and Latour (1998) gave a new method to prove stationary ergodicity based on the method given by Du and Li (1991). Moreover, Drost, Van Den Akker and Werker (2008) studied the local asymptotic normality and efficient estimation. Zhu and Joe (2006) obtained new models and results for count time series based on binomial thinning. Drost, Van Den Akker and Werker (2009) considered the semi-parametric efficient estimation for INAR(p) models. Recently, Zheng, Basawa and Datta (2006) studied the random-coefficient INAR(p) model, which is defined by the following recursive equation:

$$X_t = \sum_{i=1}^{p} \phi_i^{(t)} \circ X_{t-i} + \varepsilon_t, \qquad t \geq 1,$$

where the random parameter $\{\phi_i^{(t)}\}$ replaces the fixed $\{\phi_i\}$ values in the literature by Du and Li (1991), and is an i.i.d. sequence for fixed i and $\sum_{i=1}^{p} E(\phi_i^{(t)}) < 1$. However, in the practical-life situations, $\{\phi_i^{(t)}\}$ may be dependent in some kind of relationship. In this paper, we extend the above model to a dependence-driven random coefficient model DDRCINAR(p), where $\{\phi_{ti}\}$ is a dependence-driven sequence of random vectors with a joint distribution function $P_{\{\phi_{t1},\dots,\phi_{tp}\}}$. Therefore, this article is mainly to introduce the basic statistical properties of this model and provide some inferential methods for the relevant parameters associated with this model.

The structure of the article is as follows. In Section 2, the dependence-driven random coefficient model DDRCINAR(p) is described in detail and we show, under certain conditions, the stationarity and the ergodicity of the DDRCINAR(p) model. In Section 3, we propose three estimation methods for the DDRCINAR(p) model parameters and study their consistency and asymptotic properties. In Section 4, comparisons among the three methods for the DDRCINAR(2) model and the proportion of in-range estimates are given via simulation studies. In Section 5, two real data sets are analysed by using estimation methods in Section 3. In Section 6, we give a summary and concluding remarks. The paper ends with conditional moments used later which are provided in the Appendix.

2 The pth-order dependence-driven random coefficient integer-valued autoregressive model

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

$$X_t = \sum_{i=1}^{p} \phi_{ti} \circ X_{t-i} + \varepsilon_t, \qquad t \geq 1, \tag{2.1}$$

where $\{\varepsilon_t\}$ is an i.i.d. non-negative integer-valued sequence with a probability mass function $f_\varepsilon > 0$, such that $E(\varepsilon_t^4) < \infty$; $\{\phi_{ti}, 1 \leq i \leq p\}$ and $\{\varepsilon_t\}$ are independent each other; the joint distribution of $\{\phi_{t1}, \phi_{t2}, \dots, \phi_{tp}\}$ is given by

$$\begin{cases} p(\phi_{t1} = \phi_1, \phi_{t2} = 0, \dots, \phi_{tp} = 0) = \alpha_1; \\ p(\phi_{t1} = 0, \phi_{t2} = \phi_2, \dots, \phi_{tp} = 0) = \alpha_2; \\ \vdots \\ p(\phi_{t1} = 0, \phi_{t2} = 0, \dots, \phi_{tp} = \phi_p) = \alpha_p; \\ p(\phi_{t1} = 0, \phi_{t2} = 0, \dots, \phi_{tp} = 0) = \alpha_0, \end{cases} \tag{2.2}$$

where $\alpha_0, \alpha_1, \dots, \alpha_p$ are non-negative and $\sum_{i=0}^{p} \alpha_i = 1$. Let

$$\mu_\varepsilon = E(\varepsilon_t), \qquad \sigma_\varepsilon^2 = \mathrm{Var}(\varepsilon_t).$$

According to (2.2), we have

$$\begin{aligned} E(\phi_{ti}) &= \alpha_i \phi_i, \\ \mathrm{Cov}(\phi_{ti}, \phi_{tj}) &= -\alpha_i \alpha_j \phi_i \phi_j, \qquad \mathrm{Var}(\phi_{ti}) = \alpha_i \phi_i^2 (1 - \alpha_i). \end{aligned} \tag{2.3}$$

Now, we give some notation for the following theorem.

Let

$$\mathbf{X}_t = (X_t, X_{t-1}, \dots, X_{t-p+1})'_{1\times p},$$

$$\boldsymbol{\varepsilon}_t = (\varepsilon_t, 0, \dots, 0)'_{1\times p} \quad \text{and} \quad \boldsymbol{\mu}_\varepsilon = (\mu_\varepsilon, 0, \dots, 0)'_{1\times p}.$$

Then we have

$$\mathbf{X}_t = A_t \circ \mathbf{X}_{t-1} + \boldsymbol{\varepsilon}_t,$$

where

$$A_t = \begin{bmatrix} \phi_{t1} & \phi_{t2} & \cdots & \phi_{t,p-1} & \phi_{tp} \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix},$$

$$A = \begin{bmatrix} \alpha_1\phi_1 & \alpha_2\phi_2 & \cdots & \alpha_{p-1}\phi_{p-1} & \alpha_p\phi_p \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}.$$

Theorem 2.1. *If $\sum_{i=1}^{p} \alpha_i \phi_i < 1$ and the maximum absolute eigenvalue of $E[A_t' \otimes A_t]$ is less than 1, then there exists a unique stationary integer-valued random series $\{X_t\}$ satisfying equation (2.1). Furthermore, the process is an ergodic process.*

Proof. First, we introduce a sequence of random variables $\{Z_t^{(n)}\}_{n\in\mathcal{N}}$,

$$Z_t^{(n)} = \begin{cases} 0, & n < 0, \\ \varepsilon_t, & n = 0, \\ \sum_{i=1}^{p} \phi_{ti} \circ Z_{t-i}^{(n-i)} + \varepsilon_t, & n > 0. \end{cases} \tag{2.4}$$

Let

$$U(n,t,k) = \left|Z_t^{(n)} - Z_t^{(n-k)}\right| \quad \text{and} \quad l(n,t,k) = \min\left(Z_t^{(n)}, Z_t^{(n-k)}\right).$$

Then we have the following inequality

$$\begin{aligned} U(n,t,k) &\le \sum_{i=1}^{p} \left|\phi_{ti} \circ Z_{t-i}^{(n-i)} - \phi_{ti} \circ Z_{t-i}^{(n-i-k)}\right| \\ &= \sum_{i=1}^{p} \left| \sum_{j=1}^{Z_{t-i}^{(n-i)}} B_j^{(t,i)} - \sum_{j=1}^{Z_{t-i}^{(n-i-k)}} B_j^{(t,i)} \right| \\ &= \sum_{i=1}^{p} \sum_{j=1}^{U(n-i,t-i,k)} B_{l(n,t,k)+j}^{(t,i)} \\ &\stackrel{d}{=} \sum_{i=1}^{p} \phi_{ti} \circ U(n-i,t-i,k). \end{aligned} \tag{2.5}$$

Let

$$B = \begin{bmatrix} \alpha_1\phi_1(1-\phi_1) & \alpha_2\phi_2(1-\phi_2) & \cdots & \alpha_p\phi_p(1-\phi_p) \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix},$$

$$A_t \circ \begin{bmatrix} X_1 \\ \vdots \\ X_p \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{p} \phi_{ti} \circ X_i \\ X_1 \\ \vdots \\ X_{p-1} \end{bmatrix}.$$

Now we note that

$$\mathbf{U}_{t,k}^{(n)} = \left(U(n,t,k), U(n-1,t-1,k), \dots, U(n-p+1,t-p+1,k)\right)'.$$

Then by equation (2.4) and the above notation, following the similar argument in Latour (1998) or Zheng, Basawa and Datta (2006), we can obtain

$$\begin{aligned} E(\mathbf{U}_{t,k}^{(n)}) &\leq E(A_t \circ \mathbf{U}_{t-1,k}^{(n-1)}) = AE(\mathbf{U}_{t-1,k}^{(n-1)}) \\ &\leq \cdots \leq A^n E(\mathbf{U}_{t-n,k}^{(0)}) \\ &= A^n(\mu_\varepsilon, 0, \ldots, 0)' = A^n \boldsymbol{\mu}_\varepsilon \end{aligned}$$

and

$$E(\mathbf{U}_{t,k}^{(n)} \mathbf{U}_{t,k}^{(n)'}) \leq \operatorname{diag}(BA^{n-1}\boldsymbol{\mu}_\varepsilon) + E(A_t E(\mathbf{U}_{t-1,k}^{(n-1)} \mathbf{U}_{t-1,k}^{(n-1)'}) A_t').$$

Therefore

$$\begin{aligned} &\operatorname{vec}(E(\mathbf{U}_{t,k}^{(n)} \mathbf{U}_{t,k}^{(n)'})) \\ &\quad \leq \operatorname{vec}(\operatorname{diag}(BA^{n-1}\boldsymbol{\mu}_\varepsilon)) + \operatorname{vec}(E(A_t E(\mathbf{U}_{t-1,k}^{(n-1)} \mathbf{U}_{t-1,k}^{(n-1)'}) A_t')) \\ &\quad = \operatorname{vec}(\operatorname{diag}(BA^{n-1}\boldsymbol{\mu}_\varepsilon)) + E(A_t' \otimes A_t) \operatorname{vec}(E(\mathbf{U}_{t-1,k}^{(n-1)} \mathbf{U}_{t-1,k}^{(n-1)'})). \end{aligned} \tag{2.6}$$

Applying equation (2.5) recurrence n times, we obtain

$$\begin{aligned} \operatorname{vec}(E(\mathbf{U}_{t,k}^{(n)} \mathbf{U}_{t,k}^{(n)'})) &\leq \sum_{j=0}^{n-1} (E(A_t' \otimes A_t))^j \operatorname{vec}(\operatorname{diag}(BA^{n-j-1}\boldsymbol{\mu}_\varepsilon)) \\ &\quad + (E(A_t' \otimes A_t))^n \operatorname{vec}\left(\begin{bmatrix} \sigma_\varepsilon^2 + \mu_\varepsilon^2 & \mathbf{0}_{p-1}' \\ \mathbf{0}_{p-1} & \mathbf{0}_{p-1\times p-1} \end{bmatrix}\right). \end{aligned} \tag{2.7}$$

Because $\{A_t\}$ is an i.i.d. random matrix sequence, all the above inequalities are elementwise.

Since maximum absolute eigenvalue of $E(A_t' \otimes A_t)$ is less than 1, $[E(A_t' \otimes A_t)]^n$ converges to a null matrix. And A^n converges to a null matrix as well since $\sum_{i=1}^p \alpha_i \phi_i < 1$. Thus, using a similar method as in Latour (1998), we can prove that

$$E(\mathbf{U}_{t,k}^{(n)} \mathbf{U}_{t,k}^{(n)'}) \to \mathbf{0}, \qquad \text{as } n \to \infty.$$

The stationarity and uniqueness follow as in Du and Li (1991). □

For the ergodicity of the process, we can follow the proof in Du and Li (1991) based on Zikun (1965). The only difference between Zheng, Basawa and Datta (2006) and Du and Li (1991) is that Zheng, Basawa and Datta (2006) introduce that $\{\phi_1^{(t)}, \phi_2^{(t)}, \ldots, \phi_p^{(t)}\}$, $t \geq 1$ is an i.i.d. sequence with a cumulative distribution function P_{ϕ_i} and assume $E(\phi_i^{(t)}) = \phi_i$ $(i = 1, 2, \ldots, p)$. But we assume that the random vectors $\{\phi_{t1}, \phi_{t2}, \ldots, \phi_{tp}\}$ are dependence-driven and have the joint distribution function as (2.2). Therefore, the argument on ergodicity in Du and Li (1991) remains effective for our model. This condition, for the existence of the stationary and ergodic DDRCINAR(p) model, will define the parameter space Ω used here for the model. We also show $0 \le \alpha_i \le 1$, $0 \le \phi_i \le 1$, $i = 1, 2, \dots, p$ and $0 < \alpha_1 + \alpha_2 + \cdots + \alpha_p < 1$.

Next, we consider the problem of estimation involved in the DDRCINAR(p) model.

3 Estimation methods

We reparameterize equation (2.3) by defining

$$a_i = \alpha_i \phi_i, \qquad \sigma_{ii} = \alpha_i \phi_i^2 (1 - \alpha_i), \qquad i = 1, 2, \dots, p. \tag{3.1}$$

Let $\mathcal{F}_{t-1}$ be the σ-field generated by $X_1, X_2, \dots, X_{t-1}$, and $\mathcal{J}_{tp}$ be the σ-field generated by $\phi_{t1}, \phi_{t2}, \dots, \phi_{tp}$. Denote

$$a = (a_1, a_2, \dots, a_p)', \qquad \sigma = (\sigma_{11}, \sigma_{22}, \dots, \sigma_{pp})',$$

$$\theta = (a', \sigma', \mu_\varepsilon, \sigma_\varepsilon^2)', \qquad \vartheta = (\alpha_1, \alpha_2, \dots, \alpha_p, \phi_1, \phi_2, \dots, \phi_p, \mu_\varepsilon, \sigma_\varepsilon^2)'.$$

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

Next, we consider three different methods of parameter estimation, namely, the conditional least squares (CLS) estimators, the weighted conditional least squares (WCLS) estimators and the maximum quasi-likelihood estimators (MQE). An advantage of the three methods is that they do not require specifying the exact family for the innovations.

3.1 Conditional least squares estimators

The CLS estimates of a and μ_ε can be obtained by minimizing

$$Q_1(a, \mu_\varepsilon) = \sum_{t=p+1}^{n} u_t^2$$

with respect to a and μ_ε, where $u_t = X_t - E(X_t | \mathcal{F}_{t-1})$. This yields the estimators

$$\begin{aligned} \hat{a} = {} & \left(\sum_{t=p+1}^{n} Y_t Y_t' - \frac{1}{n-p} \sum_{t=p+1}^{n} Y_t \sum_{t=p+1}^{n} Y_t' \right)^{-1} \\ & \times \left(\sum_{t=p+1}^{n} Y_t X_t - \frac{1}{n-p} \sum_{t=p+1}^{n} Y_t \sum_{t=p+1}^{n} X_t \right), \end{aligned} \tag{3.2}$$

$$\hat{\mu}_\varepsilon = \frac{1}{n-p} \sum_{t=p+1}^{n} (X_t - Y_t' \hat{a}) \tag{3.3}$$

with $Y_t = (X_{t-1}, X_{t-2}, \dots, X_{t-p})'$.

To obtain estimates of σ and σ_ε^2, CLS is again applied to estimate the residual sequence H_t, where

$$\hat{H}_t = \left(X_t - \sum_{i=1}^{p} \hat{a}_i X_{t-i} - \hat{\mu}_\varepsilon\right)^2 + 2\sum_{j=2}^{p}\sum_{i=1}^{j-1} \hat{a}_i \hat{a}_j X_{t-i} X_{t-j} - \sum_{i=1}^{p} X_{t-i}\left(\hat{a}_i - \hat{a}_i^2\right),$$

by minimizing

$$Q_2(\sigma, \sigma_\varepsilon^2) = \sum_{t=p+1}^{n} \left(\hat{H}_t - \sum_{i=1}^{p} \sigma_{ii}\left(X_{t-i}^2 - X_{t-i}\right) - \sigma_\varepsilon^2\right)^2$$

with respect to σ and σ_ε^2. This yields the estimators

$$\begin{aligned} \hat{\sigma} = {} & \left(\sum_{t=p+1}^{n} Z_t Z_t' - \frac{1}{n-p} \sum_{t=p+1}^{n} Z_t \sum_{t=p+1}^{n} Z_t'\right)^{-1} \\ & \times \left(\sum_{t=p+1}^{n} Z_t \hat{H}_t - \frac{1}{n-p} \sum_{t=p+1}^{n} Z_t \sum_{t=p+1}^{n} \hat{H}_t\right), \end{aligned} \tag{3.4}$$

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{n-p} \sum_{t=p+1}^{n} \left(\hat{H}_t - Z_t' \hat{\sigma}\right) \tag{3.5}$$

with $Z_t = (X_{t-1}^2 - X_{t-1}, X_{t-2}^2 - X_{t-2}, \dots, X_{t-p}^2 - X_{t-p})'$.

We obtain estimates $\hat{\vartheta}$ from

$$\hat{\alpha}_i = \frac{\hat{a}_i^2}{\hat{\sigma}_{ii} + \hat{a}_i^2} \quad \text{and} \quad \hat{\phi}_i = \frac{\hat{\sigma}_{ii} + \hat{a}_i^2}{\hat{a}_i}, \qquad i = 1, 2, \dots, p. \tag{3.6}$$

The following theorem gives the strong consistency and the limited distribution of the estimates $\hat{\vartheta}$ given in equation (3.6).

Theorem 3.1. *Let $\{X_t\}$ be an* DDRCINAR(p) *process generated as in equation* (2.1) *and* (2.2) *with the conditions given in Theorem* 2.1. *Then the estimates $\hat{\vartheta}$ obtained from equation* (3.6) *will be strongly consistent and jointly asymptotically normally distributed.*

Proof. We first prove the strong consistency of $\hat{a}$, $\hat{\sigma}$, $\hat{\mu}_\varepsilon$ and $\hat{\sigma}_\varepsilon^2$. According to Theorem 2.1, $\{X_t\}_{t=1}^{\infty}$ is a stationary ergodic sequence of integrable random variables.

Let

$$g_1\left(\theta^{(1)}, \mathcal{F}_{t-1}\right) = E(X_t | \mathcal{F}_{t-1}) = \sum_{i=1}^{p} a_i X_{t-i} + \mu_\varepsilon,$$

then

$$Q_1(\theta^{(1)}) = \sum_{t=p+1}^{n} (X_t - g_1(\theta^{(1)}, \mathcal{F}_{t-1}))^2,$$

where $\theta^{(1)} = (a', \mu_\varepsilon)'$. Take a Taylor expansion of $Q_1(\theta^{(1)})$ carried out to third order terms:

$$\begin{aligned} Q_1(\theta^{(1)}) &= Q_1(\theta_0^{(1)}) + (\theta^{(1)} - \theta_0^{(1)})' \frac{\partial Q_1(\theta^{(1)})}{\partial \theta^{(1)}} \\ &\quad + \frac{1}{2}(\theta^{(1)} - \theta_0^{(1)})' V_1 (\theta^{(1)} - \theta_0^{(1)}) + R_1, \end{aligned}$$

where $V_1^{(p+1)\times(p+1)} = \frac{\partial^2 Q_1(\theta_0^{(1)})}{\partial(\theta^{(1)})^2}$ and R_1 is the usual remainder term. Obviously, it is easy to check that $g_1(\theta^{(1)}, \mathcal{F}_{t-1})$, $\frac{\partial g_1(\theta^{(1)},\mathcal{F}_{t-1})}{\partial \theta_i^{(1)}}$, $\frac{\partial^2 g_1(\theta^{(1)},\mathcal{F}_{t-1})}{\partial \theta_i^{(1)} \partial \theta_j^{(1)}}$ and $\frac{\partial^3 g_1(\theta^{(1)},\mathcal{F}_{t-1})}{\partial \theta_i^{(1)} \partial \theta_j^{(1)} \partial \theta_k^{(1)}}$ for $i, j, k \in \{1, 2, \dots, p+1\}$ satisfy all the regularity conditions in Klimko and Nelson (1978). Thus, Theorem 3.1 of Klimko and Nelson (1978) leads us to conclude that $\hat{\theta}^{(1)}$ is strongly consistent, which indicates that $\hat{a}$ and $\hat{\mu}_\varepsilon$ are strongly consistent.

Similarly, we obtain that $\hat{\sigma}$ and $\hat{\sigma}_\varepsilon^2$ are strongly consistent. Then $\hat{\vartheta}$ is strongly consistent from equation (3.6).

Next, we prove the asymptotic normality of the estimates $\hat{\vartheta}$. According to Theorem 3.1 of Hwang and Basawa (1998) or Theorem 3.1 of Nicholls and Quinn (1982), we have

$$\sqrt{n}(\hat{a} - a) \xrightarrow{d} N_p(\mathbf{0}, \Gamma^{-1} W \Gamma^{-1}), \qquad n \to +\infty,$$

where $W = E(u_t^2 Y_t Y_t') = E(\mathrm{var}(X_t|\mathcal{F}_{t-1}) Y_t Y_t')$ and $\Gamma = E(Y_t Y_t')$.

With the similar method, we can obtain

$$\begin{aligned} &\sqrt{n}(\hat{\sigma} - \sigma) \xrightarrow{d} N_p(\mathbf{0}, L^{-1}\Sigma L^{-1}), \\ &\sqrt{n}(\hat{\mu}_\varepsilon - \mu_\varepsilon) \xrightarrow{d} N(0, G), \\ &\sqrt{n}(\hat{\sigma}_\varepsilon^2 - \sigma_\varepsilon^2) \xrightarrow{d} N(0, I), \qquad n \to +\infty. \end{aligned}$$

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

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N_{2p+2}(\mathbf{0}, \Theta), \qquad n \to +\infty,$$

where

$$\Theta = \begin{bmatrix} \Gamma^{-1} W \Gamma^{-1} & \Gamma^{-1} V L^{-1} & \Gamma^{-1} M & \Gamma^{-1} D \\ L^{-1} V' \Gamma^{-1} & L^{-1} \Sigma L^{-1} & L^{-1} Q & L^{-1} B \\ M' \Gamma^{-1} & Q' L^{-1} & G & F \\ D' \Gamma^{-1} & B' L^{-1} & F & T \end{bmatrix},$$

where

$$M = E(Y_t u_t^2), \qquad V = E(u_t U_t Y_t Z_t') = E(\varphi_t Y_t Z_t'), \qquad L = E(Z_t Z_t'),$$
$$G = E(u_t^2), \qquad Q = E(Z_t U_t u_t), \qquad \Sigma = E(U_t^2 Z_t Z_t'),$$
$$D = E(Y_t u_t U_t), \qquad B = E(U_t^2 Z_t), \qquad F = E(u_t U_t), \qquad T = E(U_t^2),$$

where

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

According to Equation (3.6) and Proposition 6.4.3 of Brockwell and Davis (1987), we have

$$\sqrt{n}(\hat{\vartheta} - \vartheta) \xrightarrow{d} N_{2p+2}(\mathbf{0}, \Phi\Theta\Phi'), \qquad n \to +\infty,$$

where

$$\Phi = \begin{bmatrix} \Phi_{11} & \Phi_{12} & \Phi_{13} \\ \Phi_{21} & \Phi_{22} & \Phi_{23} \\ \Phi_{31} & \Phi_{32} & \Phi_{33} \end{bmatrix}$$

with

$$\Phi_{11} = \mathrm{diag}\left(\frac{2a_1\sigma_{11}}{(\sigma_{11}+a_1^2)^2}, \frac{2a_2\sigma_{22}}{(\sigma_{22}+a_2^2)^2}, \ldots, \frac{2a_p\sigma_{pp}}{(\sigma_{pp}+a_p^2)^2}\right),$$
$$\Phi_{12} = \mathrm{diag}\left(\frac{-a_1^2}{(\sigma_{11}+a_1^2)^2}, \frac{-a_2^2}{(\sigma_{22}+a_2^2)^2}, \ldots, \frac{-a_p^2}{(\sigma_{pp}+a_p^2)^2}\right),$$
$$\Phi_{13} = \Phi_{23} = \mathbf{0}_{p\times 2},$$
$$\Phi_{21} = \mathrm{diag}\left(\frac{a_1^2-\sigma_{11}}{a_1^2}, \frac{a_2^2-\sigma_{22}}{a_2^2}, \ldots, \frac{a_p^2-\sigma_{pp}}{a_p^2}\right),$$
$$\Phi_{22} = \mathrm{diag}\left(\frac{1}{a_1}, \frac{1}{a_2}, \ldots, \frac{1}{a_p}\right),$$
$$\Phi_{31} = \Phi_{32} = \mathbf{0}_{2\times p},$$
$$\Phi_{33} = \mathbf{I}_{2\times 2}.$$

□

3.2 Weighted conditional least squares estimators

The Conditional Least Squares (CLS) estimates, in general, are not asymptotically efficient. Because $\mathrm{Var}(X_t|\mathcal{F}_{t-1}, \mathcal{J}_{tp})$, $\mathrm{Var}(X_t|\mathcal{F}_{t-1})$, $\mathrm{Cov}(X_t, X_t^2|\mathcal{F}_{t-1}, \mathcal{J}_{tp})$, $\mathrm{Cov}(X_t, X_t^2|\mathcal{F}_{t-1})$, $\mathrm{Var}(X_t^2|\mathcal{F}_{t-1})$ and $\mathrm{Var}(X_t^2|\mathcal{F}_{t-1}, \mathcal{J}_{tp})$ depend on $X_{t-1}, X_{t-2},$

$\ldots, X_{t-p}, \phi_{t1}, \phi_{t2}, \ldots, \phi_{tp}$, we can consider WCLS estimates to improve the efficiency. In this section, we give these estimates.

Write

$$\psi_t = E\big(u_t^2|\mathcal{F}_{t-1}\big), \qquad \Psi_t = E\big(U_t^2|\mathcal{F}_{t-1}\big),$$

with U_t given in Section 3.1, then we have

$$\begin{aligned}
\psi_t &= \mathrm{var}(X_t|\mathcal{F}_{t-1}),\\
\Psi_t &= \mathrm{var}\big(X_t^2|\mathcal{F}_{t-1}\big) - 4E\big(X_t|\mathcal{F}_{t-1}\big)\,\mathrm{cov}\big(X_t, X_t^2|\mathcal{F}_{t-1}\big)\\
&\quad + 4\big(E\big(X_t|\mathcal{F}_{t-1}\big)\big)^2\,\mathrm{var}\big(X_t|\mathcal{F}_{t-1}\big).
\end{aligned}$$

We can obtain the WCLS estimates by minimizing

$$Q_3(\theta) = \sum_{t=p+1}^{n} \frac{u_t^2}{\psi_t} + \sum_{t=p+1}^{n} \frac{U_t^2}{\Psi_t}$$

with respect to θ. Since it is very difficult to derive explicit estimators of the parameters using an iterative method, we consider θ replaced with the corresponding consistent estimates by other means. In particular, we may choose to use the estimated versions of ψ_t and Ψ_t denoted by $\hat{\psi}_t$ and $\hat{\Psi}_t$ according to the CLS. Thus, we can derive $\hat{\theta}$. So with similar arguments as in Section 3.1, we can obtain the WCLS estimators of a, μ_ε, σ and σ_ε^2:

$$\begin{aligned}
\hat{a}^w &= \Bigg(\sum_{t=p+1}^{n} \frac{1}{\psi_t} Y_t Y_t' - \Bigg(\sum_{t=p+1}^{n} \frac{1}{\psi_t}\Bigg)^{-1} \sum_{t=p+1}^{n} \frac{1}{\psi_t} Y_t \sum_{t=p+1}^{n} \frac{1}{\psi_t} Y_t'\Bigg)^{-1}\\
&\quad \times \Bigg(\sum_{t=p+1}^{n} \frac{1}{\psi_t} Y_t X_t - \Bigg(\sum_{t=p+1}^{n} \frac{1}{\psi_t}\Bigg)^{-1} \sum_{t=p+1}^{n} \frac{1}{\psi_t} Y_t \sum_{t=p+1}^{n} \frac{1}{\psi_t} X_t\Bigg),
\end{aligned} \tag{3.7}$$

$$\hat{\mu}_\varepsilon^w = \Bigg(\sum_{t=p+1}^{n} \frac{1}{\psi_t}\Bigg)^{-1} \sum_{t=p+1}^{n} \frac{1}{\psi_t}\big(X_t - Y_t'\hat{a}^w\big), \tag{3.8}$$

$$\begin{aligned}
\hat{\sigma}^w &= \Bigg(\sum_{t=p+1}^{n} \frac{1}{\Psi_t} Z_t Z_t' - \Bigg(\sum_{t=p+1}^{n} \frac{1}{\Psi_t}\Bigg)^{-1} \sum_{t=p+1}^{n} \frac{1}{\Psi_t} Z_t \sum_{t=p+1}^{n} \frac{1}{\Psi_t} Z_t'\Bigg)^{-1}\\
&\quad \times \Bigg(\sum_{t=p+1}^{n} \frac{1}{\Psi_t} Z_t \hat{H}_t^w - \Bigg(\sum_{t=p+1}^{n} \frac{1}{\Psi_t}\Bigg)^{-1} \sum_{t=p+1}^{n} \frac{1}{\Psi_t} Z_t \sum_{t=p+1}^{n} \frac{1}{\Psi_t} \hat{H}_t^w\Bigg),
\end{aligned} \tag{3.9}$$

$$\hat{\sigma}_\varepsilon^{2w} = \Bigg(\sum_{t=p+1}^{n} \frac{1}{\Psi_t}\Bigg)^{-1} \sum_{t=p+1}^{n} \frac{1}{\Psi_t}\big(\hat{H}_t^w - Z_t'\hat{\sigma}^w\big), \tag{3.10}$$

where Y_t and Z_t are given in Section 3.1, and

$$\hat{H}_t^w = \left(X_t - \sum_{i=1}^{p} \hat{a}_i^w X_{t-i} - \hat{\mu}_\varepsilon^w\right)^2 + 2\sum_{j=2}^{p}\sum_{i=1}^{j-1} \hat{a}_i^w \hat{a}_j^w X_{t-i} X_{t-j}$$
$$- \sum_{i=1}^{p} X_{t-i}(\hat{a}_i^w - (\hat{a}_i^w)^2).$$

Thus, we can also obtain estimators $\hat{\vartheta}^w$ by using the similar equations given in equation (3.6). These estimator given in (3.7), (3.8), (3.9) and (3.10) are strongly consistent from the ergodic theorem. However, when the sample size n is small, the mean squared errors (MSE) are large, a high proportion of estimates fall outside Ω in simulations and the above estimators cannot be guaranteed to be positive estimators for σ and σ_ε^2, the results of which are presented in Section 4.

Remark 3.1. The proof of the strongly consistence of $\hat{\vartheta}^w$ is omitted here since it is difficult relatively.

Remark 3.2. The small sample size n is usually less than 1000, which is presented in Section 4.

3.3 Maximum quasi-likelihood estimators

The MQEs for the DDRCINAR(p) model can be based on the p-dimensional stochastic process $\{X_t, X_t^2, \ldots, X_t^p\}$. The resulting system of estimating equations is given by:

$$\sum_{t=p+1}^{n} B_1 \times B_2 \times B_3 = 0, \tag{3.11}$$

where

$$B_1 = \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1,2p+2} \\ e_{21} & e_{22} & \cdots & e_{2,2p+2} \\ \vdots & \vdots & \ddots & \vdots \\ e_{p1} & e_{p2} & \cdots & e_{p,2p+2} \end{bmatrix}',$$

$$B_2 = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1p} \\ v_{21} & v_{22} & \cdots & v_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ v_{p1} & v_{p2} & \cdots & v_{pp} \end{bmatrix}^{-1},$$

$$B_3 = (X_t - E(X_t|\mathcal{F}_{t-1}), X_t^2 - E(X_t^2|\mathcal{F}_{t-1}), \ldots, X_t^p - E(X_t^p|\mathcal{F}_{t-1}))',$$

where

$$e_{ij} = \frac{\partial E(X_t^i|\mathcal{F}_{t-1})}{\partial \alpha_j}, \qquad e_{i,j+p} = \frac{\partial E(X_t^i|\mathcal{F}_{t-1})}{\partial \phi_j},$$

$$e_{i,2p+1} = \frac{\partial E(X_t^i|\mathcal{F}_{t-1})}{\partial \mu_\varepsilon}, \qquad e_{i,2p+2} = \frac{\partial E(X_t^i|\mathcal{F}_{t-1})}{\partial \sigma_\varepsilon^2},$$

$$v_{ij} = \operatorname{cov}(X_t^i, X_t^j|\mathcal{F}_{t-1}) = E(X_t^{i+j}|\mathcal{F}_{t-1}) - E(X_t^i|\mathcal{F}_{t-1})E(X_t^j|\mathcal{F}_{t-1}),$$
$$i, j = 1, 2, \dots, p.$$

This nonlinear system of equations can be solved using an iterative method to obtain the MQEs $\hat{\vartheta}$ of parameter vector ϑ and the conditional moments used in equation (3.11) are given in the Appendix. Hutton and Nelson (1986) gave regularity conditions for the existence, strong consistency and asymptotic normality of the MQEs and showed that they are optimal in Godambe's sense. However, for the DDRCINAR(p) model of order $p \geq 2$, it is difficult to prove these regularity conditions because of the complexity of the algebraic expressions given in the appendix. Next, we consider some properties of MQEs when $p = 1$.

Let $\xi = (\sigma_{11}, \eta, \sigma_\varepsilon^2)'$, where $\eta = a_1(1 - a_1) - \sigma_{11}$ and $\beta = (a_1, \mu_\varepsilon)'$. Thus, the expression for one-step conditional variance

$$V_\xi(X_t|X_{t-1}) = v_{11} = \sigma_{11}X_{t-1}^2 + \eta X_{t-1} + \sigma_\varepsilon^2.$$

According to (3.11) and (3.1), a set of MQEs estimating equations take the form:

$$\begin{cases} \sum_{t=2}^{n} V_\xi^{-1}(X_t|X_{t-1})(X_t - a_1X_{t-1} - \mu_\varepsilon) = 0, \\ \sum_{t=2}^{n} V_\xi^{-1}(X_t|X_{t-1})X_{t-1}(X_t - a_1X_{t-1} - \mu_\varepsilon) = 0. \end{cases} \tag{3.12}$$

Note that the presence of ξ in the expression for the conditional variance makes the corresponding estimating equations complicated and intractable in the general case. Therefore, we propose substituting a suitable consistent estimator $\hat{\xi}$ of ξ obtained by other means and then solve the estimators of (3.12). This approach leads to the following closed form estimator of β:

$$\begin{pmatrix} \tilde{a}_1 \\ \tilde{\mu}_\varepsilon \end{pmatrix} = \begin{pmatrix} \sum_{t=2}^{n} X_{t-1}V_{\hat{\xi}}^{-1}(X_t|X_{t-1}) & \sum_{t=2}^{n} V_{\hat{\xi}}^{-1}(X_t|X_{t-1}) \\ \sum_{t=2}^{n} X_{t-1}^2V_{\hat{\xi}}^{-1}(X_t|X_{t-1}) & \sum_{t=2}^{n} X_{t-1}V_{\hat{\xi}}^{-1}(X_t|X_{t-1}) \end{pmatrix}^{-1} \times \begin{pmatrix} \sum_{t=2}^{n} X_tV_{\hat{\xi}}^{-1}(X_t|X_{t-1}) \\ \sum_{t=2}^{n} X_tX_{t-1}V_{\hat{\xi}}^{-1}(X_t|X_{t-1}) \end{pmatrix}. \tag{3.13}$$

A consistent estimator of ξ is proposed next that can be used in (3.13).

Proposition 3.1. *Let X_t be a* DDRCINAR(1) *model, then the following estimators are consistent*:

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{n}\sum_{t=2}^{n}(X_t - \hat{a}_1 X_{t-1} - \hat{\mu}_\varepsilon)^2 - \frac{\hat{\sigma}_{11}}{n}\sum_{t=2}^{n}(X_{t-1}^2 - X_{t-1}) - \frac{\hat{a}_1 - \hat{a}_1^2}{n}\sum_{t=2}^{n} X_{t-1},$$

$$\hat{\eta} = \hat{a}_1 - \hat{a}_1^2 - \hat{\sigma}_{11},$$

where $\hat{a}_1$, $\hat{\sigma}_{11}$ and $\hat{\mu}_\varepsilon$ are consistent estimators of a_1, σ_{11} and μ_ε. In practice, we can use the CLS or WCLS estimators of a_1, σ_{11} and μ_ε.

Proof. Let $A_n = \frac{1}{n}\sum_{t=2}^{n}(X_t - a_1 X_{t-1} - \mu_\varepsilon)^2$, $B_n = \frac{1}{n}\sum_{t=2}^{n} X_{t-1}^2$ and $C_n = \frac{1}{n}\sum_{t=2}^{n} X_{t-1}$. By Theorem 1.1 of Billingsley (1961), $A_n \xrightarrow{a.s.} E((X_t - a_1 X_{t-1} - \mu_\varepsilon)^2) = \sigma_\varepsilon^2 + \sigma_{11}(\gamma_1 - \gamma_2) + (a_1 - a_1^2)\gamma_2$, $B_n \xrightarrow{a.s.} \gamma_1$ and $C_n \xrightarrow{a.s.} \gamma_2$, where $\gamma_1 = E(X_\infty^2)$, $\gamma_2 = E(X_\infty)$ and X_∞ denotes the limiting random variable corresponding to the stationary of the process. Therefore,

$$\begin{aligned}
\hat{\sigma}_\varepsilon^2 &= A_n - A_n + \frac{1}{n}\sum_{t=2}^{n}(X_t - \hat{a}_1 X_{t-1} - \hat{\mu}_\varepsilon)^2 - \frac{\hat{\sigma}_{11}}{n}\sum_{t=2}^{n}(X_{t-1}^2 - X_{t-1}) \\
&\quad - \frac{\hat{a}_1 - \hat{a}_1^2}{n}\sum_{t=2}^{n} X_{t-1} \\
&= A_n + (\hat{a}_1 - a_1)\big((\hat{a}_1 + a_1 - 2)B_n + 2\hat{\mu}_\varepsilon C_n\big) \\
&\quad + (\hat{\mu}_\varepsilon - \mu_\varepsilon)\big(\hat{\mu}_\varepsilon + \mu_\varepsilon - 2(1 + a_1)C_n\big) - \hat{\sigma}_{11}(B_n - C_n) - \big(\hat{a}_1 - \hat{a}_1^2\big)C_n \\
&\xrightarrow{p} \sigma_\varepsilon^2.
\end{aligned}$$

Similar arguments lead to $\hat{\eta} \xrightarrow{p} \eta$. □

Remark 3.3. $\hat{\alpha}_1$ and $\hat{\phi}_1$ are also consistent by Proposition 3.1 and (3.6).

Asymptotic normality of the MQEs estimators in (3.13) is established in the following theorem.

Theorem 3.2. *The joint limit distribution of the MQEs estimators $(\tilde{a}_1, \tilde{\mu}_\varepsilon)$ given by* (3.13) *is*

$$\sqrt{n}\begin{pmatrix}\tilde{a}_1 - a_1 \\ \tilde{\mu}_\varepsilon - \mu_\varepsilon\end{pmatrix} \xrightarrow{d} \mathrm{N}\big(\mathbf{0}, T^{-1}(\xi)Q(\xi)T^{-1}(\xi)\big), \qquad n \to +\infty,$$

where

$$Q(\xi)=\begin{pmatrix} T_1(\xi) & T_3(\xi) \\ T_3(\xi) & T_2(\xi) \end{pmatrix},$$

$$T^{-1}(\xi)=\big(T_3^2(\xi)-T_1(\xi)T_2(\xi)\big)^{-1}\begin{pmatrix} T_3(\xi) & -T_1(\xi) \\ -T_2(\xi) & T_3(\xi) \end{pmatrix},$$

where $T_1(\xi)=E[V_\xi^{-1}(X_2|X_1)]$, $T_2(\xi)=E[X_1^2V_\xi^{-1}(X_2|X_1)]$ *and* $T_3(\xi)=E[X_1V_\xi^{-1}(X_2|X_1)]$.

Proof. First, we suppose ξ is known. For the following estimation equations:

$$S_n^{(1)}(\xi,\beta)=\sum_{t=2}^{n}V_\xi^{-1}(X_t|X_{t-1})(X_t-a_1X_{t-1}-\mu_\varepsilon),$$

$$S_n^{(2)}(\xi,\beta)=\sum_{t=2}^{n}V_\xi^{-1}(X_t|X_{t-1})X_{t-1}(X_t-a_1X_{t-1}-\mu_\varepsilon),$$

we have

$$\begin{aligned}&E\big[V_\xi^{-1}(X_t|X_{t-1})(X_t-a_1X_{t-1}-\mu_\varepsilon)|\mathcal{F}_{t-1}\big]\\&\quad=V_\xi^{-1}(X_t|X_{t-1})E\big[(X_t-a_1X_{t-1}-\mu_\varepsilon)|\mathcal{F}_{t-1}\big]=0\end{aligned}$$

and

$$E\big[S_t^{(1)}(\xi,\beta)|\mathcal{F}_{t-1}\big]=S_{t-1}^{(1)}(\xi,\beta).$$

Thus, $\{S_t^{(1)}(\xi,\beta),\mathcal{F}_t,t\geq 0\}$ is a martingale. By Theorem 1.1 of Billingsley (1961),

$$\begin{aligned}&\frac{1}{n}\sum_{t=2}^{n}V_\xi^{-2}(X_t|X_{t-1})(X_t-a_1X_{t-1}-\mu_\varepsilon)^2\\&\quad\xrightarrow{a.s.}E\big[V_\xi^{-2}(X_2|X_1)(X_2-a_1X_1-\mu_\varepsilon)^2\big]\\&\quad=E\big[E\big(V_\xi^{-2}(X_2|X_1)(X_2-a_1X_1-\mu_\varepsilon)^2|X_1\big)\big]=E\big[V_\xi^{-1}(X_2|X_1)\big]=T_1(\xi).\end{aligned}$$

By Corollary 3.2 of Hall and Heyde (1980), the martingale MQEs applies and we get

$$\frac{1}{\sqrt{n}}S_n^{(1)}(\xi,\beta)\xrightarrow{d}N\big(0,T_1(\xi)\big),\qquad n\to+\infty.$$

Similarly,

$$\begin{aligned}
&\frac{1}{n}\sum_{t=2}^{n} V_\xi^{-2}(X_t|X_{t-1})X_{t-1}^2(X_t - a_1X_{t-1} - \mu_\varepsilon)^2 \\
&\quad \xrightarrow{a.s.} E\big[V_\xi^{-2}(X_2|X_1)X_1^2(X_2 - a_1X_1 - \mu_\varepsilon)^2\big] \\
&\quad = E\big[E\big(V_\xi^{-2}(X_2|X_1)X_1^2(X_2 - a_1X_1 - \mu_\varepsilon)^2|X_1\big)\big] \\
&\quad = E\big[X_1^2 V_\xi^{-1}(X_2|X_1)\big] = T_2(\xi)
\end{aligned}$$

and

$$\frac{1}{\sqrt{n}} S_n^{(2)}(\xi,\beta) \xrightarrow{d} N\big(0, T_2(\xi)\big), \qquad n \to +\infty.$$

Again by Cramer–Wold device, for any $c = (c_1, c_2)'$, c_1 and $c_2 \in \mathcal{R}$ are not both 0. When $n \to +\infty$, we have

$$\frac{c'}{\sqrt{n}}\begin{pmatrix} S_n^{(1)}(\xi,\beta) \\ S_n^{(2)}(\xi,\beta)\end{pmatrix} \xrightarrow{d} N\big(0, E\big[V_\xi^{-2}(X_2|X_1)(c_2X_1 + c_1)^2(X_2 - a_1X_1 - \mu_\varepsilon)^2\big]\big),$$

implying

$$\frac{1}{\sqrt{n}}\begin{pmatrix} S_n^{(1)}(\xi,\beta) \\ S_n^{(2)}(\xi,\beta)\end{pmatrix} \xrightarrow{d} N\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix} T_1(\xi) & T_3(\xi) \\ T_3(\xi) & T_2(\xi)\end{pmatrix}\right), \qquad n \to +\infty, \tag{3.14}$$

where $T_3(\xi) = E[V_\xi^{-2}(X_2|X_1)X_1(X_2 - a_1X_1 - \mu_\varepsilon)^2] = E[X_1 V_\xi^{-1}(X_2|X_1)]$.

Now, we replace $V_\xi^{-2}(X_t|X_{t-1})$ by $V_{\hat\xi}^{-2}(X_t|X_{t-1})$, where $\hat\xi$ is a consistent estimator of ξ. Then we want

$$\frac{1}{\sqrt{n}}\begin{pmatrix} S_n^{(1)}(\hat\xi,\beta) \\ S_n^{(2)}(\hat\xi,\beta)\end{pmatrix} \xrightarrow{d} N\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix} T_1(\xi) & T_3(\xi) \\ T_3(\xi) & T_2(\xi)\end{pmatrix}\right), \qquad n \to +\infty. \tag{3.15}$$

To obtain this we need to prove that

$$\frac{1}{\sqrt{n}} S_n^{(i)}(\hat\xi,\beta) - \frac{1}{\sqrt{n}} S_n^{(i)}(\xi,\beta) \xrightarrow{p} 0, \qquad i = 1, 2. \tag{3.16}$$

Let $R_n(\xi) = (1/\sqrt{n})S_n^{(1)}(\xi,\beta)$. Then $\forall\epsilon > 0$ and $\delta > 0$ such that $\xi - \delta\mathbf{1} > 0$, where $\mathbf{1}$ is the unit vector, we have

$$\begin{aligned}
&P\big(\big|R_n(\hat\xi) - R_n(\xi)\big| > \epsilon\big) \\
&\quad \le P\big(|\hat\sigma_{11} - \sigma_{11}| > \delta\big) + p\big(|\hat\eta - \eta| > \delta\big) \\
&\qquad + p\big(|\hat\sigma_\varepsilon^2 - \sigma_\varepsilon^2| > \delta\big) \\
&\qquad + P\Big(\sup_{\{|\sigma_1 - \sigma_{11}|<\delta, |\eta_1 - \eta|<\delta, |\sigma_2^2 - \sigma_\varepsilon^2|<\delta\}} \big|R_n(\xi_1) - R_n(\xi)\big| > \epsilon\Big),
\end{aligned}$$

where $\xi_1 = (\sigma_1, \eta_1, \sigma_2^2)'$. Let $D = \{|\sigma_1 - \sigma_{11}| < \delta, |\eta_1 - \eta| < \delta, |\sigma_2^2 - \sigma_\varepsilon^2| < \delta\}$. If $\hat{\xi}$ is a consistent estimator of ξ, then we just need to prove that

$$P\Big(\sup_D \big|R_n(\xi_1) - R_n(\xi)\big| > \epsilon\Big) \xrightarrow{p} 0.$$

By Markov inequality,

$$\begin{aligned}
&P\Big(\sup_D \big|R_n(\xi_1) - R_n(\xi)\big| > \epsilon\Big) \\
&\quad\le \frac{1}{\epsilon^2} E\Big(\sup_D \big(R_n(\xi_1) - R_n(\xi)\big)^2\Big) \\
&\quad= \frac{1}{\epsilon^2} E\Bigg(\sup_D \frac{1}{n}\sum_{t=2}^n \big(V_{\xi_1}^{-1}(X_t|X_{t-1}) - V_\xi^{-1}(X_t|X_{t-1})\big)^2 \\
&\qquad \times (X_t - a_1 X_{t-1} - \mu_\varepsilon)^2\Bigg) \\
&\quad= \frac{1}{\epsilon^2} E\Big(\sup_D \big(V_{\xi_1}^{-1}(X_2|X_1) - V_\xi^{-1}(X_2|X_1)\big)^2 (X_2 - a_1 X_1 - \mu_\xi)^2\Big) \\
&\quad= \frac{1}{\epsilon^2} E\Bigg(\sup_D \frac{((\sigma_1 - \sigma_{11})X_1^2 + (\eta_1 - \eta)X_1 + (\sigma_2^2 - \sigma_\varepsilon^2))^2}{V_{\xi_1}^2(X_2|X_1) V_\xi^2(X_2|X_1)} \\
&\qquad \times (X_2 - a_1 X_1 - \mu_\varepsilon)^2\Bigg) \\
&\quad= \frac{1}{\epsilon^2} E\Bigg(\sup_D \frac{((\sigma_1 - \sigma_{11})X_1^2 + (\eta_1 - \eta)X_1 + (\sigma_2^2 - \sigma_\varepsilon^2))^2}{V_{\xi_1}^2(X_2|X_1) V_\xi(X_2|X_1)}\Bigg) \\
&\quad\le \frac{1}{\epsilon^2} \sup_D \big\{(\sigma_1 - \sigma_{11})^2 c_1 + (\eta_1 - \eta)^2 c_2 \\
&\qquad + \big(\sigma_2^2 - \sigma_\varepsilon^2\big)^2 c_3 + 2c_4\big|(\sigma_1 - \sigma_{11})(\eta_1 - \eta)\big| \\
&\qquad + 2c_5\big|(\sigma_1 - \sigma_{11})\big(\sigma_2^2 - \sigma_\varepsilon^2\big)\big| + 2c_6\big|(\eta_1 - \eta)\big(\sigma_2^2 - \sigma_\varepsilon^2\big)\big|\big\} \\
&\quad\le \frac{C\delta^2}{\epsilon^2},
\end{aligned}$$

where $\{c_i, i = 1, \dots, 6\}$ are finite moments and C is a positive constant. Similar argument can be used for $\frac{1}{\sqrt{n}} S_n^{(2)}(\xi, \beta)$. When δ goes to zero, we get our assertion which in turn establishes (3.15). Similarly, we have

$$\frac{1}{n}\sum_{t=2}^n V_\xi^{-1}(X_t|X_{t-1}) - \frac{1}{n}\sum_{t=2}^n V_{\hat{\xi}}^{-1}(X_t|X_{t-1}) \xrightarrow{p} 0,$$

$$\frac{1}{n}\sum_{t=2}^{n} X_{t-1}V_{\xi}^{-1}(X_t|X_{t-1}) - \frac{1}{n}\sum_{t=2}^{n} X_{t-1}V_{\hat{\xi}}^{-1}(X_t|X_{t-1}) \xrightarrow{p} 0,$$

$$\frac{1}{n}\sum_{t=2}^{n} X_{t-1}^2 V_{\xi}^{-1}(X_t|X_{t-1}) - \frac{1}{n}\sum_{t=2}^{n} X_{t-1}^2 V_{\hat{\xi}}^{-1}(X_t|X_{t-1}) \xrightarrow{p} 0.$$

Therefore, by the above and Theorem 1.1 of Billingsley (1961), we have

$$\begin{aligned}(A_1 - A_2)^{-1} \times A_3 &\xrightarrow{p} \bigl(T_3^2(\xi) - T_1(\xi)T_2(\xi)\bigr)^{-1} \begin{pmatrix} T_3(\xi) & -T_1(\xi) \\ -T_2(\xi) & T_3(\xi) \end{pmatrix} \\ &= T^{-1}(\xi),\end{aligned} \tag{3.17}$$

where

$$A_1 = \left(\frac{1}{n}\sum_{t=2}^{n} X_{t-1}V_{\hat{\xi}}^{-1}(X_t|X_{t-1})\right)^2,$$

$$A_2 = \left(\frac{1}{n}\sum_{t=2}^{n} V_{\hat{\xi}}^{-1}(X_t|X_{t-1})\right)\left(\frac{1}{n}\sum_{t=2}^{n} X_{t-1}^2 V_{\hat{\xi}}^{-1}(X_t|X_{t-1})\right)$$

$$A_3 = \begin{pmatrix} \frac{1}{n}\sum_{t=2}^{n} X_{t-1}V_{\hat{\xi}}^{-1}(X_t|X_{t-1}) & -\frac{1}{n}\sum_{t=2}^{n} V_{\hat{\xi}}^{-1}(X_t|X_{t-1}) \\ -\frac{1}{n}\sum_{t=2}^{n} X_{t-1}^2 V_{\hat{\xi}}^{-1}(X_t|X_{t-1}) & \frac{1}{n}\sum_{t=2}^{n} X_{t-1}V_{\hat{\xi}}^{-1}(X_t|X_{t-1}) \end{pmatrix}.$$

After some algebra, we have

$$\begin{pmatrix} \tilde{a}_1 - a_1 \\ \tilde{\mu}_\varepsilon - \mu_\varepsilon \end{pmatrix} = n^{-1}(A_1 - A_2)^{-1} \times A_3 \times \begin{pmatrix} S_n^{(1)}(\hat{\xi}, \beta) \\ S_n^{(2)}(\hat{\xi}, \beta) \end{pmatrix}.$$

Therefore, by (3.15) and (3.17),

$$\sqrt{n}\begin{pmatrix} \tilde{a}_1 - a_1 \\ \tilde{\mu}_\varepsilon - \mu_\varepsilon \end{pmatrix} \xrightarrow{d} N\bigl(\mathbf{0}, T^{-1}(\xi)' Q(\xi) T^{-1}(\xi)\bigr), \qquad n \to +\infty,$$

where

$$Q(\xi) = \begin{pmatrix} T_1(\xi) & T_3(\xi) \\ T_3(\xi) & T_2(\xi) \end{pmatrix}. \qquad \square$$

Remark 3.4. When $p = 1$, the MQEs is equivalent to modified quasi-likelihood (MQL) estimation method proposed by Zheng, Basawa and Datta (2007). Thus, the proof of the above theorem can be also obtained by Zheng, Basawa and Datta (2007).

4 Simulations

Consider the model 2.1 with $p = 2$, that is, DDRCINAR(2) model, where ε_t is an i.i.d. poisson sequence with mean λ, and

$$E(A_t' \otimes A_t) = \begin{bmatrix} \alpha_1\phi_1^2 & 0 & \alpha_1\phi_1 & \alpha_2\phi_2 \\ \alpha_1\phi_1 & 0 & 1 & 0 \\ 0 & \alpha_2\phi_2^2 & 0 & 0 \\ \alpha_2\phi_2 & 0 & 0 & 0 \end{bmatrix}.$$

By Theorem 2.1, the conditions for the existence of this stationary and ergodic DDRCINAR(2) model are: $0 \le \alpha_1, \alpha_2 \le 1$; $0 < \alpha_1 + \alpha_2 < 1$; $0 \le \phi_1, \phi_2 \le 1$; $\alpha_1\phi_1 + \alpha_2\phi_2 < 1$ and maximum absolute eigenvalue of $E(A_t' \otimes A_t) < 1$. These conditions define the parameter space Ω used here. A simulation study was conducted by generating DDRCINAR(2) processes, each of which are from four samples of parameter values for α_i, ϕ_i and λ, namely

(a) sample 1: $\alpha_1 = 0.4$, $\alpha_2 = 0.5$, $\phi_1 = 0.6$, $\phi_2 = 0.7$ and $\lambda = 1$,
(b) sample 2: $\alpha_1 = 0.4$, $\alpha_2 = 0.5$, $\phi_1 = 0.6$, $\phi_2 = 0.7$ and $\lambda = 2$,
(c) sample 3: $\alpha_1 = 0.25$, $\alpha_2 = 0.25$, $\phi_1 = 0.5$, $\phi_2 = 0.5$ and $\lambda = 1$,
(d) sample 4: $\alpha_1 = 0.25$, $\alpha_2 = 0.25$, $\phi_1 = 0.5$, $\phi_2 = 0.5$ and $\lambda = 2$.

Figure 1 on page 656 gives four typical sample pathes for a sample size 200 about the DDRCINAR(2) model. We use R for the random number generation

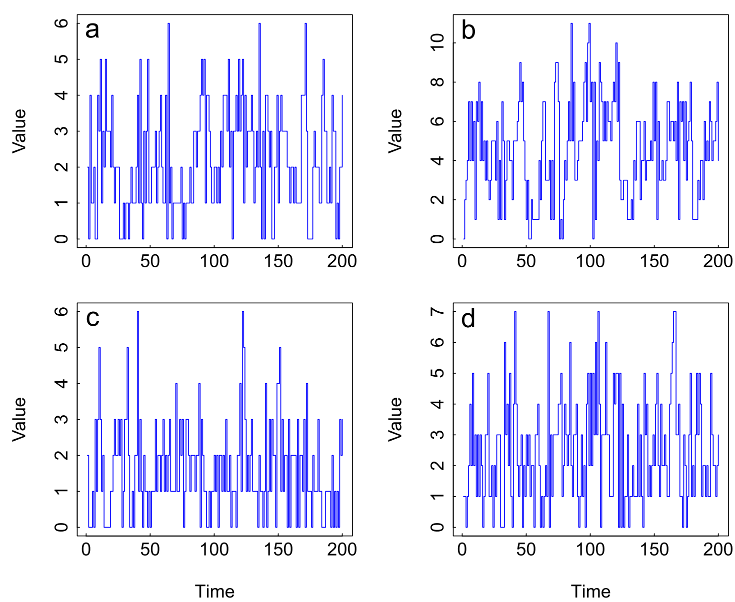

Figure 1 *Samples* 1, 2, 3 *and* 4 *paths of Model.*

Table 1 *Mean values of three sets of estimates for samples* 1 *and* 2

n	α_1	α_2	ϕ_1	ϕ_2	λ
$\alpha_1 = 0.4, \alpha_2 = 0.5, \phi_1 = 0.6, \phi_2 = 0.7, \lambda = 1$					
CLS					
10,000	0.4006	0.5001	0.5985	0.7004	1.0000
30,000	0.3990	0.5009	0.6022	0.6995	0.9988
WCLS					
10,000	0.3999	0.5002	0.5987	0.7005	1.0003
30,000	0.3990	0.5007	0.6019	0.6998	0.9990
MQE					
10,000	0.3999	0.4997	0.6026	0.7041	0.9910
30,000	0.3991	0.5003	0.6036	0.7014	0.9949
$\alpha_1 = 0.4, \alpha_2 = 0.5, \phi_1 = 0.6, \phi_2 = 0.7, \lambda = 2$					
CLS					
10,000	0.4010	0.5004	0.5990	0.7006	1.9986
30,000	0.4008	0.4997	0.5998	0.7004	1.9991
WCLS					
10,000	0.4011	0.5002	0.5990	0.7011	1.9974
30,000	0.4004	0.4997	0.6001	0.7004	1.9992
MQE					
10,000	0.4014	0.5001	0.6016	0.7035	1.9842
30,000	0.4005	0.4998	0.6019	0.7021	1.9899

and sample sizes $n = 200, 500, 1000, 10{,}000, 30{,}000$ and 500 replications were used. For WCLS and MQE estimates, ψ_t, Ψ_t and v_{ij} are estimated by using CLS. We use the mean squared errors (MSE), that is,

$$\frac{1}{n}\sum_{j=1}^{n}(\vartheta^{\text{est}} - \vartheta)^2, \tag{4.1}$$

to evaluate the performance of the estimators, where n is the number of realizations and ϑ^{est} denotes any estimator of ϑ. Table 1 on page 657 enumerates the estimates of parameters for samples 1 and 2, with similar results given for samples 3 and 4 in Table 3 on page 659. The representative results about MSE and the percent lying in Ω for samples 1, 2, 3 and 4 are summarized in Tables 2 on page 658 and 4 on page 660, respectively.

Values outside the allowed range for ϑ might easily be obtained for small sample sizes of $n = 200, 500$ and 1000. By (3.6), the estimates of α_i and ϕ_i are decided by a and σ. Therefore, when the sample size is very small, there emerge negative estimated values of α_i and ϕ_i. We thus need to adjust negative estimates in a somewhat ad hoc manner. In Tables 3 on page 659 and 4 on page 660, such estimated

Table 2 *MSE and per cent within parenthese space of three sets of estimates for samples* 1 *and* 2

n	α_1	α_2	ϕ_1	ϕ_2	λ	% in Ω
$\alpha_1 = 0.4, \alpha_2 = 0.5, \phi_1 = 0.6, \phi_2 = 0.7, \lambda = 1$						
CLS						
10,000	0.0005	0.0004	0.0007	0.0004	0.0008	100.00
30,000	0.0002	0.0001	0.0002	0.0001	0.0003	100.00
WCLS						
10,000	0.0004	0.0003	0.0005	0.0004	0.0007	100.00
30,000	0.0001	0.0001	0.0002	0.0001	0.0002	100.00
MQE						
10,000	0.0004	0.0003	0.0004	0.0003	0.0002	100.00
30,000	0.0001	0.0001	0.0002	0.0001	0.0001	100.00
$\alpha_1 = 0.4, \alpha_2 = 0.5, \phi_1 = 0.6, \phi_2 = 0.7, \lambda = 2$						
CLS						
10,000	0.0004	0.0003	0.0004	0.0002	0.0023	100.00
30,000	0.0001	0.0001	0.0001	0.0001	0.0010	100.00
WCLS						
10,000	0.0003	0.0002	0.0003	0.0002	0.0020	100.00
30,000	0.0001	0.0001	0.0001	0.0001	0.0008	100.00
MQE						
10,000	0.0003	0.0002	0.0002	0.0001	0.0008	100.00
30,000	0.0001	0.0001	0.0001	0.0000	0.0003	100.00

have been adjusted by taking account that the restrictions on α_i and ϕ_i imply,

$$\begin{cases} 0 < a_1 + a_2 < 1, \\ \sigma_{11}\sigma_{22} > a_1^2 a_2^2, \\ 0 < \sigma_{ii} < 0.25, \\ \sigma_{ii} < a_i(1 - a_i), \qquad i = 1, 2. \end{cases} \tag{4.2}$$

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

As the sample size increases, three sets of estimates seem to converge to the true parameter values, indicating consistency. However, the MQEs seem fall within Ω with a higher proportion of times than the existing estimators of CLS and WCLS, which indicate an improvement. Observing Tables 2 on page 658 and 4 on page 660, it is conclusion that MQEs dominate CLS and WCLS in terms of the MSE, which is accordance with our expectation. Meanwhile, when the sample size is increased, it is better to estimate the parameters for these three estimation methods, which shows large sample size may be needed to obtain reasonable results.

In order to more fully compare the three sets of estimators in terms of the percentage of estimates that fall within Ω, we calculate and examine the Ω of the two

Table 3 *Mean values of three sets of estimates for samples* 3 *and* 4

n	α_1	α_2	ϕ_1	ϕ_2	λ
$\alpha_1=\alpha_2=0.25$, $\phi_1=\phi_2=0.5$, $\lambda=1$					
CLS					
200	0.4032	0.4280	0.3991	0.4010	0.9530
500	0.3486	0.3797	0.4132	0.4292	0.9655
1000	0.3115	0.3235	0.4550	0.4581	0.9793
10,000	0.2535	0.2601	0.4981	0.4908	1.0005
30,000	0.2510	0.2515	0.5004	0.4998	0.9999
WCLS					
200	0.4022	0.4364	0.3998	0.3964	0.9565
500	0.3522	0.3839	0.4268	0.4293	0.9659
1000	0.3125	0.3249	0.4551	0.4582	0.9804
10,000	0.2526	0.2569	0.4973	0.4947	1.0004
30,000	0.2507	0.2507	0.5005	0.5014	0.9999
MQE					
200	0.3721	0.4174	0.4014	0.4095	0.9508
500	0.3313	0.3548	0.4473	0.4466	0.9675
1000	0.2968	0.3341	0.4560	0.4729	0.9815
10,000	0.2551	0.2598	0.4992	0.4957	0.9932
30,000	0.2528	0.2524	0.5004	0.5013	0.9955
$\alpha_1=\alpha_2=0.25$, $\phi_1=\phi_2=0.5$, $\lambda=2$					
CLS					
200	0.3101	0.3321	0.3836	0.4431	2.0713
500	0.2854	0.2859	0.4680	0.4717	2.0168
1000	0.2667	0.2697	0.4759	0.4853	2.0172
10,000	0.2529	0.2510	0.4975	0.4991	2.0025
30,000	0.2528	0.2492	0.4967	0.5016	2.0003
WCLS					
200	0.3204	0.3054	0.4304	0.4594	2.0745
500	0.2879	0.2705	0.4511	0.4866	2.0162
1000	0.2636	0.2663	0.4772	0.4884	2.0152
10,000	0.2523	0.2501	0.4982	0.4994	2.0031
30,000	0.2521	0.2498	0.4976	0.5008	2.0001
MQE					
200	0.3541	0.3571	0.4419	0.4469	1.8986
500	0.3100	0.2907	0.4576	0.4845	1.9422
1000	0.2763	0.2776	0.4798	0.4899	1.9623
10,000	0.2562	0.2547	0.4999	0.4999	1.9881
30,000	0.2545	0.2518	0.4981	0.5021	1.9918

estimators for a range of different parameter values when $\phi_1=\phi_2=0.6$, $\lambda=1$, and $\phi_1=\phi_2=0.6$, $\lambda=2$. The sum of α_1 and α_2 is confined within the range of [0.0,1.0]. And, for each of these two parameters, different values range from 0.0 to 1.0, on a grid of 0.10. All possible samples of α_1 and α_2 are examined.

Table 4 *MSE and per cent within parenthese space of three sets of estimates for samples* 3 *and* 4

n	α_1	α_2	ϕ_1	ϕ_2	λ	% in Ω
$\alpha_1 = \alpha_2 = 0.25$, $\phi_1 = \phi_2 = 0.5$, $\lambda = 1$						
CLS						
200	0.0910	0.1042	0.0799	0.0764	0.0079	56.60
500	0.0546	0.0662	0.0456	0.0489	0.0036	76.00
1000	0.0264	0.0356	0.0280	0.0297	0.0015	92.80
10,000	0.0016	0.0018	0.0053	0.0048	0.0004	100.00
30,000	0.0005	0.0005	0.0014	0.0014	0.0001	100.00
WCLS						
200	0.0922	0.1078	0.0802	0.0776	0.0071	55.20
500	0.0556	0.0686	0.0467	0.0505	0.0035	75.00
1000	0.0262	0.0355	0.0287	0.0309	0.0014	92.80
10,000	0.0013	0.0014	0.0035	0.0036	0.0004	100.00
30,000	0.0004	0.0004	0.0010	0.0011	0.0001	100.00
MQE						
200	0.0772	0.0983	0.0668	0.0685	0.0084	57.60
500	0.0506	0.0571	0.0403	0.0436	0.0033	79.40
1000	0.0212	0.0211	0.0225	0.0243	0.0013	94.40
10,000	0.0012	0.0013	0.0029	0.0030	0.0002	100.00
30,000	0.0004	0.0004	0.0009	0.0010	0.0001	100.00
$\alpha_1 = \alpha_2 = 0.25$, $\phi_1 = \phi_2 = 0.5$, $\lambda = 2$						
CLS						
200	0.0865	0.0926	0.0643	0.0705	0.0795	62.20
500	0.0457	0.0312	0.0431	0.0353	0.0327	92.60
1000	0.0163	0.0165	0.0193	0.0207	0.0160	98.40
10,000	0.0010	0.0009	0.0021	0.0019	0.0014	100.00
30,000	0.0003	0.0003	0.0006	0.0006	0.0005	100.00
WCLS						
200	0.0866	0.0902	0.0714	0.0615	0.0797	64.80
500	0.0430	0.0243	0.0359	0.0300	0.0291	93.80
1000	0.0141	0.0146	0.0163	0.0176	0.0137	98.60
10,000	0.0010	0.0009	0.0017	0.0015	0.0012	100.00
30,000	0.0003	0.0003	0.0005	0.0005	0.0004	100.00
MQE						
200	0.0622	0.0678	0.0464	0.0504	0.0609	78.80
500	0.0255	0.0129	0.0224	0.0210	0.0109	97.20
1000	0.0114	0.0119	0.0121	0.0137	0.0050	99.40
10,000	0.0010	0.0008	0.0016	0.0014	0.0005	100.00
30,000	0.0003	0.0002	0.0004	0.0005	0.0002	100.00

Tables 5 on page 661 and 6 on page 662 show the proportion of estimates for the parameters, $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\phi}_1$, $\hat{\phi}_2$ and $\hat{\lambda}$. Obviously, when α_1 or α_2 is 0 or the sum of α_1 and α_2 is 1, the percentage of in-range estimates is very small, which indicates that the problem of out-of-range estimates is severe near the boundaries of Ω.

Table 5 *Percentage of in-range estimates for $\phi_1 = \phi_2 = 0.6$ and $\lambda = 1$*

	α_1										
α_2	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
CLS											
0.00	1.8	5.8	10.8	12.2	14.4	16.4	18.6	17.4	18.2	14.2	4.4
0.10	6.4	24.4	39.6	47.2	49.0	50.0	55.4	51.0	42.4	18.8	
0.20	13.4	42.6	65.4	74.8	78.6	79.0	74.0	70.2	26.0		
0.30	15.4	51.8	75.6	86.4	90.2	88.8	75.4	32.4			
0.40	17.2	54.8	82.2	89.6	90.4	76.6	33.0				
0.50	18.0	57.8	83.0	90.6	79.4	32.4					
0.60	23.2	59.4	82.6	79.6	30.6						
0.70	19.4	61.8	71.6	33.8							
0.80	17.2	49.2	29.4								
0.90	16.6	17.8									
1.00	7.4										
WCLS											
0.00	1.2	3.6	10.8	13.0	13.8	15.0	18.6	16.8	17.0	16.0	6.0
0.10	6.8	25.0	39.6	49.0	53.0	53.8	59.6	55.2	46.0	21.0	
0.20	11.4	42.0	65.6	76.4	80.4	84.6	80.6	72.2	31.8		
0.30	12.8	53.4	76.8	90.0	92.4	92.4	80.4	37.4			
0.40	17.6	56.6	84.6	91.8	92.8	80.0	40.4				
0.50	17.6	59.8	88.2	91.0	81.8	37.8					
0.60	20.6	61.6	84.0	83.6	38.0						
0.70	19.4	64.4	78.4	41.6							
0.80	20.2	54.8	38.0								
0.90	17.4	27.2									
1.00	9.4										
MQE											
0.00	51.6	58.4	62.6	68.6	70.4	73.8	77.8	79.0	72.6	65.2	41.0
0.10	55.8	66.4	76.0	82.6	84.2	84.2	85.6	82.6	75.4	43.4	
0.20	67.2	78.0	90.6	91.2	93.6	95.2	90.2	85.8	50.0		
0.30	72.6	82.6	93.6	95.4	96.8	95.2	86.2	52.4			
0.40	77.0	84.0	96.0	96.6	96.4	86.2	58.8				
0.50	77.2	87.0	96.4	95.6	87.8	56.2					
0.60	74.2	89.0	92.2	88.8	52.6						
0.70	73.8	86.0	86.6	53.6							
0.80	71.2	81.4	52.0								
0.90	64.6	48.8									
1.00	36.6										

Generally, it is true that the three estimation methods improve when α_1 or α_2 or both increase(s). Specifically, it can be seen, from Tables 5 on page 661 and 6 on page 662, that MQEs perform better than the estimators of CLS and WCLS in terms of the proportion of within-Ω estimates.

Table 6 *Percentage of in-range estimates for* $\phi_1 = \phi_2 = 0.6$ *and* $\lambda = 2$

	α_1										
α_2	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
CLS											
0.00	2.0	10.4	16.4	16.6	19.2	21.4	21.4	21.6	21.0	16.0	6.6
0.10	11.6	34.6	53.8	59.2	62.4	64.4	66.6	64.8	60.6	22.6	
0.20	18.0	55.8	79.0	84.4	88.6	91.2	90.6	83.6	30.0		
0.30	19.6	64.0	86.0	95.2	97.2	97.6	95.2	33.0			
0.40	22.6	69.2	90.8	98.0	99.0	97.2	36.0				
0.50	24.0	70.2	92.2	98.6	97.2	37.2					
0.60	25.8	69.0	93.6	97.0	32.2						
0.70	27.8	72.2	91.0	37.4							
0.80	25.0	71.6	32.8								
0.90	22.8	26.6									
1.00	10.0										
WCLS											
0.00	2.0	8.4	17.4	18.2	19.4	19.6	20.0	21.0	21.0	16.4	8.8
0.10	11.2	34.2	51.0	63.4	63.0	67.8	67.6	67.4	67.2	28.8	
0.20	16.0	56.2	79.0	86.2	90.4	93.0	92.6	89.6	36.8		
0.30	19.4	65.0	87.0	96.6	98.4	99.0	96.4	41.2			
0.40	23.4	69.0	92.2	99.0	99.4	97.6	41.6				
0.50	24.6	73.4	94.6	99.4	98.6	45.2					
0.60	24.6	73.0	97.0	96.6	42.8						
0.70	28.0	76.4	94.4	45.6							
0.80	26.0	74.2	41.2								
0.90	22.8	33.8									
1.00	12.0										
MQE											
0.00	59.6	66.4	76.8	78.2	78.6	78.8	77.4	77.0	76.0	73.4	41.8
0.10	67.0	83.8	88.8	91.2	94.4	93.0	90.6	91.6	89.6	49.4	
0.20	78.4	91.6	98.0	98.8	99.2	99.8	97.6	96.2	53.6		
0.30	79.2	94.4	99.0	99.8	99.8	99.6	98.4	58.0			
0.40	79.8	96.0	99.4	100.0	99.8	99.0	61.6				
0.50	78.2	95.2	99.6	100.0	99.4	61.8					
0.60	78.0	94.4	99.2	99.0	57.2						
0.70	76.8	94.0	96.2	59.6							
0.80	76.6	93.0	57.8								
0.90	72.0	49.6									
1.00	37.4										

We also simulated other representative parameter samples. We find that good estimates and high proportion of within-Ω estimates can be derived when α_1 and α_2 are both large, and poor estimates and low proportion of within-Ω estimates are derived when α_1 or α_2 is small.

5 Real data analysis

In this section, we will show how the model and methods from Section 3 can be applied to two real data time series. Moreover, we will consider two kinds of criteria to compare different models to the real data sets. The first kind is the root mean squared (RMS) errors defined by

$$\mathrm{RMS} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \hat{X}_i)^2}, \tag{5.1}$$

and the second is defined as

$$\mathrm{MSE} = \frac{1}{m}\sum_{i=1}^{m}(X_{n-m+i} - \hat{X}_{n-m+i})^2, \tag{5.2}$$

where $\hat{X}_k = E(X_k|\mathcal{F}_{t-1})$ and take $m = 30$. The MSE criteria is studied by Li, Lian and Zhu (2016). The predictive performance of models is evaluated according to the two criteria.

5.1 Epileptic seizure counts analysis

Franke and Seligmann (1993), Latour (1998) and Zheng, Basawa and Datta (2006) analysed these data by using different methods. Franke and Seligmann (1993) used conditional maximum likelihood method for their SINAR(1) model, Latour (1998) used conditional least squares method for the fixed coefficient INAR(p) model, and Zheng, Basawa and Datta (2006) used conditional least squares method and modified quasi-likelihood for the random coefficient INAR(p) model. We will reconsider the first half the time series of patient number 2, corresponding to the period before the patient had submitted to medical treatment.

The data were extracted from Figure 22.3 of Franke and Seligmann (1993) and are presented in Figure 2 on page 664. Note that the counts vary from 0 to 5, the sample mean equals 0.6612 and the variance 0.8592. Moreover, the plots of ACF and PACF are given in Figure 3 on page 664. From the graphs, we would assume that the model for the process is a dependence-driven random coefficient integer-valued autoregressive model as follows:

$$X_t = \phi_{t6} \circ X_{t-6} + \phi_{t,14} \circ X_{t-14} + \varepsilon_t, \tag{5.3}$$

where the joint distribution of $\{\phi_{t6}, \phi_{t,14}\}$ is given by

$$\begin{cases} p(\phi_{t6} = \phi_6, \phi_{t,14} = 0) = \alpha_6; \\ p(\phi_{t6} = 0, \phi_{t,14} = \phi_{14}) = \alpha_{14}; \\ p(\phi_{t6} = 0, \phi_{t,14} = 0) = 1 - \alpha_6 - \alpha_{14}. \end{cases} \tag{5.4}$$

Using the methods proposed in Section 3, the following results are given in Table 7 on page 665.

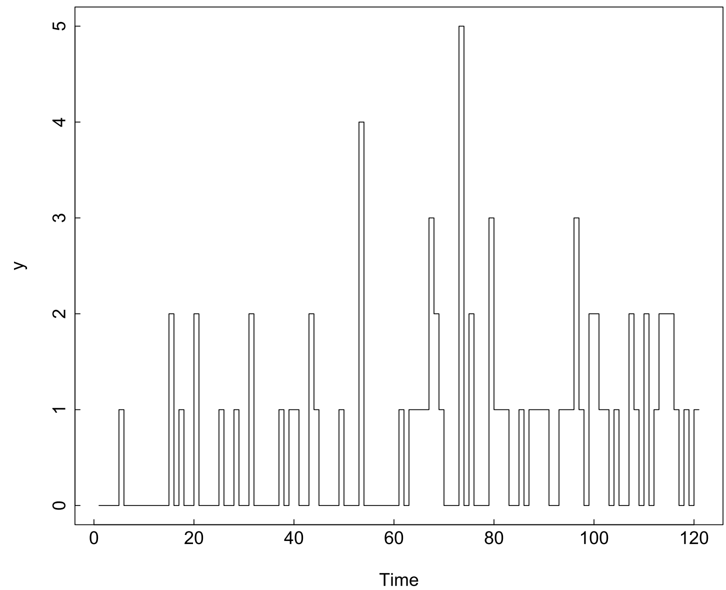

Figure 2 *Seizure counts plot*.

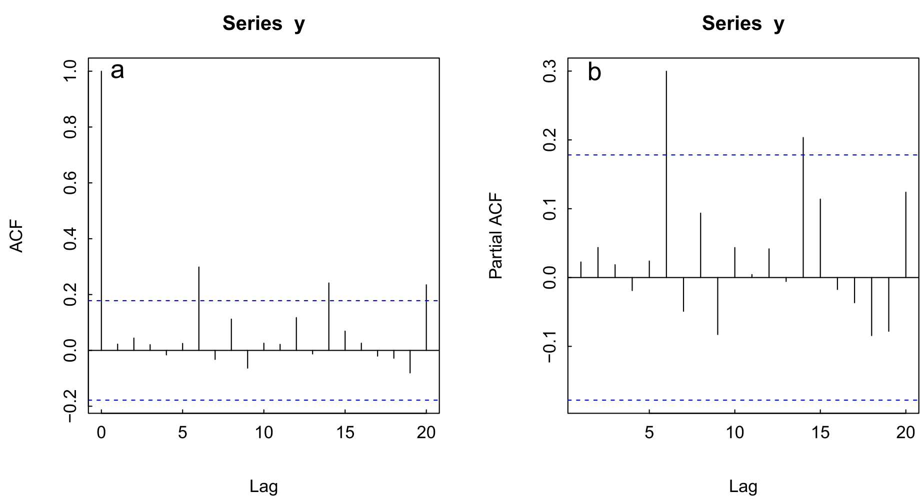

Figure 3 *Seizure counts ACF and PACF plots*.

Since $\hat{\alpha}_{14} > 1$ for CLS and WCLS methods, and $\hat{\alpha}_6 + \hat{\alpha}_{14} > 1$ for MQE in Table 7 on page 665, we should treat $\phi_{t,14}$ as the fixed coefficient, i.e. take $\sigma_{14} = 0$. Then, using WCLS method, we obtain the variance estimation of σ_ε^2 is that $\hat{\sigma}_\varepsilon^2 = 0.5015$. The difference between $\hat{\sigma}_\varepsilon^2$ and $\hat{\mu}_\varepsilon$ is 0.0603, which indicates it is reasonable that we assume the error ε_t has Poisson distribution with expectation

Table 7 *Parameter estimation for model* (5.3)

Method	Parameter				
	α_6	α_{14}	ϕ_6	ϕ_{14}	μ_ε
CLS	0.3664	1.0423	0.7001	0.2342	0.4079
WCLS	0.2397	1.3373	0.8756	0.1843	0.4412
MQE	0.2468	0.8390	0.8946	0.2962	0.4421

0.4421. Thus, our model for these data is degenerated to

$$X_t = \phi_{t6} \circ X_{t-6} + \phi_{14} \circ X_{t-14} + \varepsilon_t, \tag{5.5}$$

where the distribution of $\{\phi_{t6}\}$ is given by

$$\begin{cases} p(\phi_{t6} = \phi_6) = \alpha_6; \\ p(\phi_{t6} = 0) = 1 - \alpha_6. \end{cases} \tag{5.6}$$

According to the above assumptions, maximum quasi-likelihood estimators for Model (5.5) are given by $\check{\alpha}_6 = 0.2459$, $\check{\phi}_6 = 0.8942$, $\check{\phi}_{14} = 0.2455$ and $\check{\mu}_\varepsilon = 0.4446$. Then by (5.1) and (5.2), we have

$$\text{RMS} = 0.8837,$$

$$\text{MSE} = 0.5878.$$

If we use the fixed coefficient model, where ε_t is poisson-distributed, then we have the following results for the MQE method:

$$\tilde{\phi}_6 = 0.2277, \qquad \tilde{\phi}_{14} = 0.2271, \qquad \tilde{\mu}_\varepsilon = 0.4614,$$

similarly, we have

$$\text{RMS} = 0.8839,$$

$$\text{MSE} = 0.5919.$$

The fitting results are summarized in Figure 4 on page 666. Figure 4 on page 666 shows the standardized residuals, the histograms of standardized residuals, ACF and PACF plots of residuals under two Models. As is known in Figure 4 on page 666, the residuals are stationary series. Furthermore, the residual mean and variance of model (5.5) and fixed-coefficient model are $(-0.0119, 0.7882)$ and $(-0.0228, 0.7882)$, respectively, which show that the model (5.5) is closest to a normal distribution relatively.

From the above results, we can see that the random coefficient model has the smallest RMS and MSE. On the one hand, based on the RMS and MSE alone, one may prefer to select the coefficient model for these data. On the other hand, one may prefer to select the coefficient model for these data because of the innovation that the autoregressive parameters are dependence-driven random variables with a joint distribution function.

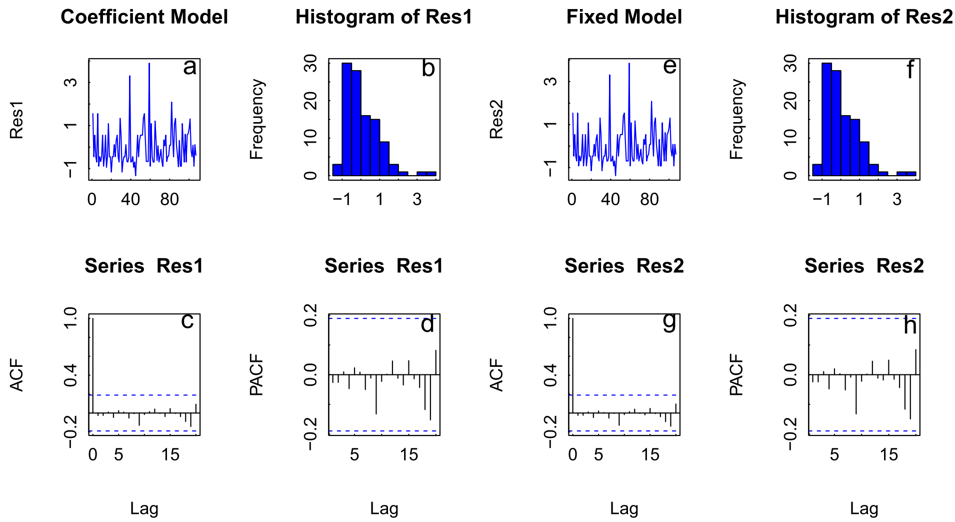

Figure 4 *Diagnostic checking plots under different models for the monthly Seizure data*. (a), (e) *Standardized residuals*; (b), (f) *Histograms of standardized residuals*; (c), (g) *ACF plots of residuals*; (d), (h) *PACF plots of residuals*.

Remark 5.1. The confidence internals in Figure 3 on page 664 are

$$(-2/\sqrt{n}, 2/\sqrt{n}) = (-0.1818, 0.1818),$$

which can be obtained by Cryer and Chan (2008). The confidence internals in the following ACF and PACF plots can be derived by the same way.

5.2 Precinct rape counts analysis

The data are obtained from the rape data section of the Forecasting Principles site (http://www.forecastingprinciples.com). There are 132 observations, starting in January 1991 and ending in December 2001. Note that the counts vary from 0 to 9. The sample mean and variance are 2.2348 and 2.9902, respectively. The plots of the time series, its ACF and PACF are given in Figure 5 on page 667. Analyzing the diagrams we conclude that the first-order autoregressive model is appropriate for the given data series. Therefore, we consider two models for the data. They are:

Model I.

$$X_t = \phi_{t,10} \circ X_{t-10} + \varepsilon_t, \tag{5.7}$$

the distribution of $\{\phi_{t,10}\}$ is given by

$$\begin{cases} p(\phi_{t,10} = \phi_{10}) = \alpha_{10}; \\ p(\phi_{t,10} = 0) = 1 - \alpha_{10}. \end{cases}$$

Where $\{\varepsilon_t\}$ is an i.i.d. poisson sequence with mean λ and ϕ_{10} is in [0, 1).

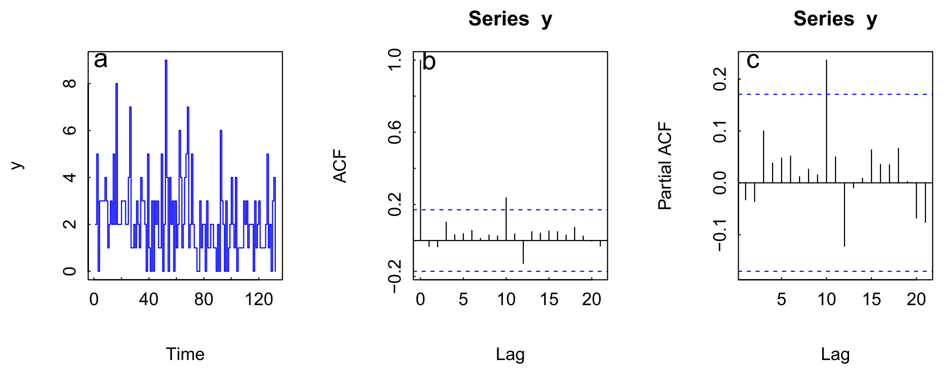

Figure 5 *Precinct rape counts sample, ACF and PACF plots.*

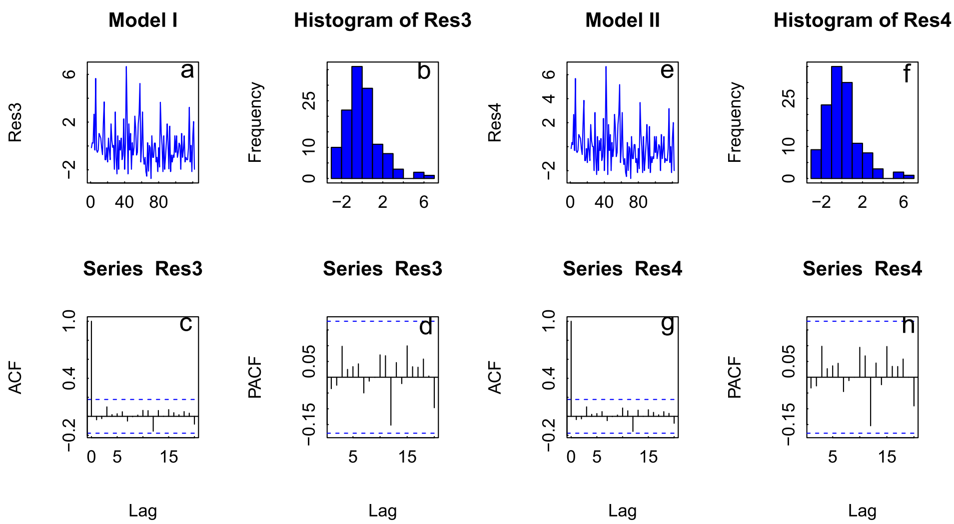

Figure 6 *Diagnostic checking plots under different models for the monthly Precinct rape data.* (a), (e) *Standardized residuals*; (b), (f) *Histograms of standardized residuals*; (c), (g) *ACF plots of residuals*; (d), (h) *PACF plots of residuals.*

Model II.

$$X_t = \varphi_{10} \circ X_{t-10} + \varepsilon_t$$

where $\{\varepsilon_t\}$ is an i.i.d. poisson sequence with mean μ and φ_{10} is fixed in $[0, 1)$.

The fitting results are summarized in Figure 6 on page 667 and Table 8 on page 668. Figure 6 on page 667 shows the standardized residuals, the histograms of standardized residuals, ACF and PACF plots of residuals under two models. As is known in Figure 6 on page 667, the residuals are stationary series. From the histograms of Figure 6 on page 667, the two models are well closed to a normal

Table 8 *Parameters, RMS and MSE*

	MQE	CLS	WCLS
Model I	$\hat{\alpha}_{10}^{\mathrm{MQE}} = 0.3709$	$\hat{\alpha}_{10}^{\mathrm{CLS}} = 0.5753$	$\hat{\alpha}_{10}^{\mathrm{WCLS}} = 0.4079$
	$\hat{\phi}_{10}^{\mathrm{MQE}} = 0.5198$	$\hat{\phi}_{10}^{\mathrm{CLS}} = 0.4408$	$\hat{\phi}_{10}^{\mathrm{WCLS}} = 0.5425$
	$\hat{\lambda}_{10}^{\mathrm{MQE}} = 1.7614$	$\hat{\lambda}_{10}^{\mathrm{CLS}} = 1.6231$	$\hat{\lambda}_{10}^{\mathrm{WCLS}} = 1.6874$
RMS	1.6916		
MSE	1.7195		
Model II	$\hat{\varphi}_{10}^{\mathrm{MQE}} = 0.1645$	$\hat{\varphi}_{10}^{\mathrm{CLS}} = 0.2536$	$\hat{\varphi}_{10}^{\mathrm{WCLS}} = 0.2341$
	$\hat{\mu}_{10}^{\mathrm{MQE}} = 1.8287$	$\hat{\mu}_{10}^{\mathrm{CLS}} = 1.6231$	$\hat{\mu}_{10}^{\mathrm{WCLS}} = 1.6672$
RMS	1.6954		
MSE	1.7210		

distribution. In Table 8 on page 668, we also give the predicted values RMS and MSE for each model. Moreover, The intuitionistic reason that the CLS estimators of λ and μ are relatively small is as follows:

It is easy to obtain the CLS estimations,

$$\begin{aligned}\hat{a}^{\mathrm{CLS}} = {} & \left(\sum_{t=11}^{n} Y_t Y_t' - \frac{1}{n-10}\sum_{t=11}^{n} Y_t \sum_{t=11}^{n} Y_t'\right)^{-1} \\ & \times \left(\sum_{t=11}^{n} Y_t X_t - \frac{1}{n-10}\sum_{t=11}^{n} Y_t \sum_{t=11}^{n} X_t\right),\end{aligned} \tag{5.8}$$

$$\hat{\lambda}^{\mathrm{CLS}} = \frac{1}{n-10}\sum_{t=11}^{n}\left(X_t - Y_t' \hat{a^{\mathrm{CLS}}}\right), \tag{5.9}$$

where $n = 132$, the data which we use to estimate λ are $X_{11}, X_{12}, \ldots, X_{132}$ without $X_1, X_2, \ldots, X_{10}$, thus some important "information" are lost. Then a similar argument can be applied to WCLS. But MQEs are based on the 10-dimensional stochastic process $\{X_t, X_t^2, \ldots, X_t^{10}\}$, where the value of t is from 11 to 132. Thus, MQEs may derive more "information". Therefore, we recommend MQE method to analyse the real data. As can be seen, RMS and MSE are smaller for the Model I than Model II. For this data, all criteria show that Model I performs best.

6 Summary and conclusion

In this paper, we have introduced a pth-order dependence-driven random coefficient integer-valued autoregressive model for count data. The autoregressive coefficient is allowed to vary randomly over time. The stationarity and ergodicity of the process are established. MQE, CLS and WCLS methods are used to estimate the parameters. Some of their asymptotic properties are obtained.

In the simulation study, we have shown three estimation methods for the DDRCINAR(p) model. And we conclude that a very large sample size may be needed to obtain reasonable results. Without considering the time factor, we recommend to use MQE to estimate the parameters in the DDRCINAR(p) model. We also consider the proportion of in-range parameter estimates of the DDRCINAR(p) model. It is concluded that the MQE method performs better than CLS and WCLS methods on the part of the proportion of within-Ω estimates, especially near the boundaries of the parameter space. The model is applied to two real data sets. It is shown that the dependence-driven random coefficient model (DDRCINAR(p)) is suitable for the real data sets.

Appendix

In Section 2, we introduce our model DDRCINAR(p). Next, we give the conditional moments used in Section 4. From equation (2.1) we have

$$\begin{aligned} X_t^m &= \left(\sum_{i=1}^{p} \phi_i^{(t)} \circ X_{t-i} + \varepsilon_t\right)^m \\ &= \sum_{k=0}^{m-1} C_m^k \sum_{i=1}^{p} (\phi_i^{(t)} \circ X_{t-i})^{m-k} \varepsilon_t^k + \varepsilon_t^m + R_m, \qquad m = 1, 2, \dots, 2p, \end{aligned} \tag{A.1}$$

where C_m^k is the number of combinations of size k from $1, 2, \dots, m$ and

$$R_m = X_t^m - \sum_{k=0}^{m-1} C_m^k \sum_{i=1}^{p} (\phi_i^{(t)} \circ X_{t-i})^{m-k} \varepsilon_t^k - \varepsilon_t^m.$$

According to $\{\varepsilon_s, s \geq t\}$ and $\{\phi_i^{(s)}, i = 1, 2, \dots, p; s \geq t\}$ are independent of $\{X_s, s \leq t-1\}$, we have $E(R_m|\mathcal{F}_{t-1}) = 0$ from (2.1) and (2.2). Then

$$\begin{aligned} &E(X_t^m|\mathcal{F}_{t-1}, \mathcal{J}_{tp}) \\ &\quad = E\left(\left[\sum_{k=0}^{m-1} C_m^k \sum_{i=1}^{p} (\phi_i^{(t)} \circ X_{t-i})^{m-k} \varepsilon_t^k + \varepsilon_t^m\right] \Big| \mathcal{F}_{t-1}, \mathcal{J}_{tp}\right) \\ &\quad = \sum_{k=0}^{m-1} C_m^k \sum_{i=1}^{p} E\left((\phi_i^{(t)} \circ X_{t-i})^{m-k} | \mathcal{F}_{t-1}, \mathcal{J}_{tp}\right) E(\varepsilon_t^k) + E(\varepsilon_t^m). \end{aligned} \tag{A.2}$$

Let $S_i = \phi_i^{(t)} \circ X_{t-i}$, we know that S_i is a conditional binomial distribution. Denote

$$S_i|\mathcal{F}_{t-1}, \mathcal{J}_{tp} \sim B(X_{t-i}, \phi_i^{(t)}).$$

Then the conditional expected recursion formula of S_i is as follows

$$\begin{aligned} E\left(S_i^{k+1}|\mathcal{F}_{t-1},\mathcal{J}_{tp}\right) = \phi_i^{(t)}\left(1-\phi_i^{(t)}\right)\frac{dE(S_i^k|\mathcal{F}_{t-1},\mathcal{J}_{tp})}{d\phi_i^{(t)}} \\ + X_{t-i}\phi_i^{(t)}E\left(S_i^k|\mathcal{F}_{t-1},\mathcal{J}_{tp}\right). \end{aligned} \tag{A.3}$$

Since ε_t has a poisson distribution with parameter λ, the quantities $E(\varepsilon_t^k)$ can be obtained by the expected recursion formula

$$E\left(\varepsilon_t^{k+1}\right) = \lambda\frac{dE(\varepsilon_t^k)}{d\lambda} + \lambda E\left(\varepsilon_t^k\right), \qquad k = 1, 2, \ldots, m. \tag{A.4}$$

Taking conditional expectations on both sides of (A.1), we obtain the following formula as follows

$$E\left(X_t^m|\mathcal{F}_{t-1}\right) = \sum_{k=0}^{m-1} C_m^k \sum_{i=1}^{p} E\left[E\left(S_i^{m-k}|\mathcal{F}_{t-1},\mathcal{J}_{tp}\right)|\mathcal{F}_{t-1}\right]E\left(\varepsilon_t^k\right) + E\left(\varepsilon_t^m\right). \tag{A.5}$$

Next, substituting (A.3) and (A.4) into equation (A.5), we can obtain $E(X_t^m|\mathcal{F}_{t-1})$, $m = 1, 2, \ldots, 2p$.

Specifically, we have

$$E(X_t|\mathcal{F}_{t-1}) = \sum_{i=1}^{p}\alpha_i\phi_i X_{t-i} + \lambda,$$

$$\begin{aligned} \mathrm{Var}(X_t|\mathcal{F}_{t-1}) &= \sum_{i=1}^{p}\left[\alpha_i\phi_i(1-\phi_i)X_{t-i} + \alpha_i\phi_i^2X_{t-i}^2\right] + 2\lambda\sum_{i=1}^{p}\alpha_i\phi_i X_{t-i} \\ &\quad + \lambda + \lambda^2 - \left(\sum_{i=1}^{p}\alpha_i\phi_i X_{t-i} + \lambda\right)^2, \end{aligned}$$

$$\begin{aligned} &\mathrm{Cov}\left(X_t, X_t^2|\mathcal{F}_{t-1}\right) \\ &\quad = \sum_{i=1}^{p}\left[\alpha_i\phi_i(1-\phi_i)(1-2\phi_i)X_{t-i} + 3\alpha_i\phi_i^2(1-\phi_i)X_{t-i}^2 + \alpha_i\phi_i^3X_{t-i}^3\right] \\ &\qquad + 3\lambda\sum_{i=1}^{p}\left[\alpha_i\phi_i(1-\phi_i)X_{t-i} + \alpha_i\phi_i^2X_{t-i}^2\right] \\ &\qquad + 3\left(\lambda+\lambda^2\right)\sum_{i=1}^{p}\alpha_i\phi_i X_{t-i} + \lambda + 3\lambda^2 + \lambda^3 \end{aligned}$$

$$
\begin{aligned}
&- \left(\sum_{i=1}^{p} \alpha_i \phi_i X_{t-i} + \lambda\right)\left(\sum_{i=1}^{p} [\alpha_i \phi_i (1-\phi_i) X_{t-i} + \alpha_i \phi_i^2 X_{t-i}^2]\right) \\
&- \left(\sum_{i=1}^{p} \alpha_i \phi_i X_{t-i} + \lambda\right)\left(2\lambda \sum_{i=1}^{p} \alpha_i \phi_i X_{t-i} + \lambda + \lambda^2\right),
\end{aligned}
$$

$$
\begin{aligned}
&\mathrm{Var}\big(X_t^2 | \mathcal{F}_{t-1}\big) \\
&= \sum_{i=1}^{p} [\alpha_i \phi_i (1-\phi_i)\big(1 - 6\phi_i + 6\phi_i^2\big) X_{t-i} + \alpha_i \phi_i^2 (1-\phi_i)(7 - 11\phi_i) X_{t-i}^2] \\
&+ \sum_{i=1}^{p} [6\alpha_i \phi_i^3 (1-\phi_i) X_{t-i}^3 + \alpha_i \phi_i^4 X_{t-i}^4] \\
&+ 4\lambda \sum_{i=1}^{p} [\alpha_i \phi_i (1-\phi_i)(1-2\phi_i) X_{t-i} + 3\alpha_i \phi_i^2 (1-\phi_i) X_{t-i}^2 + \alpha_i \phi_i^3 X_{t-i}^3] \\
&+ 6\big(\lambda + \lambda^2\big) \sum_{i=1}^{p} [\alpha_i \phi_i (1-\phi_i) X_{t-i} + \alpha_i \phi_i^2 X_{t-i}^2] \\
&+ 4\big(\lambda + 3\lambda^2 + \lambda^3\big) \sum_{i=1}^{p} \alpha_i \phi_i X_{t-i} + \lambda + 7\lambda^2 + 6\lambda^3 + \lambda^4 \\
&- \left(\sum_{i=1}^{p} [\alpha_i \phi_i (1-\phi_i) X_{t-i} + \alpha_i \phi_i^2 X_{t-i}^2] + 2\lambda \sum_{i=1}^{p} \alpha_i \phi_i X_{t-i} + \lambda + \lambda^2\right)^2.
\end{aligned}
$$

Acknowledgments

This work is supported by National Natural Science Foundation of China (No. 11731015, 11571051, 11501241), Jilin Province Natural Science Foundation (No. 20150520053JH, 20170101057JC), Program for Changbaishan Scholars of Jilin Province (2015010), and Science and Technology Program of Jilin Educational Department during the "13th Five-Year" Plan Period (No.2016-399). Thank you all for helping me writing this LaTeX sample file. Meanwhile, the authors are very grateful to the referees for their valuable suggestions and comments which have greatly improved this paper.

References

Al-Osh, M. A. and Alzaid, A. A. (1987). First-order integer-valued autoregressive (INAR(1)) process. *Journal of Time Series Analysis* **8**, 261–275. MR0903755

Alzaid, A. and Al-Osh, M. (1988). First-order integer-valued autoregressive (INAR(1)) process: distributional and regression properties. *Statistica Neerlandica* **42**, 53–61. MR0959714
Alzaid, A. A. and Al-Osh, M. (1990). An integer-valued pth-order autoregressive structure (INAR(p)) process. *Journal of Applied Probability* **27**, 314–324. MR1052303
Billingsley, P. (1961). *Statistical Inference for Markov Processes*. Chicago: University of Chicago Press. MR0123419
Brännäs, K. and Hellström, J. (2001). Generalized integer-valued autoregression. *Econometric Reviews* **20**, 425–443. MR1868341
Brockwell, P. J. and Davis, R. A. (1987). *Time Series: Theory and Methods*. Media: Springer. MR0868859
Cryer, J. D. and Chan, K. S. (2008). Time series analysis: With applications in R.
Davis, R. A., Dunsmuir, W. T. M. and Wang, Y. (1999). Modeling time series of count data. *Statistics Textbooks and Monographs* **158**, 63–113. MR1724696
Drost, F. C., Van Den Akker, R. and Werker, B. J. (2008). Local asymptotic normality and efficient estimation for INAR(p) models. *Journal of Time Series Analysis* **29**, 783–801. MR2450896
Drost, F. C., Van Den Akker, R. and Werker, B. J. (2009). Efficient estimation of auto-regression parameters and innovation distributions for semiparametric integer-valued AR(p) models. *Journal of the Royal Statistical Society, Series B, Statistical Methodology* **71**, 467–485. MR2649605
Du, J. G. and Li, Y. (1991). The integer-valued autoregressive (INAR(p)) model. *Journal of Time Series Analysis* **12**, 129–142. MR1108796
Franke, J. and Seligmann, T. (1993). Conditional maximum likelihood estimates for INAR(1) processes and their application to modelling epileptic seizure counts. *Developments in Time Series Analysis*, 310–330. MR1292273
Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theory and Its Applications*. New York: Academic Press. MR0624435
Hutton, J. E. and Nelson, P. I. (1986). Quasi-likelihood estimation for semimartingales. *Stochastic Processes and Their Applications* **22**, 245–257. MR0860935
Hwang, S. Y. and Basawa, I. V. (1998). Parameter estimation for generalized random coefficient autoregressive processes. *Journal of Statistical Planning and Inference* **68**, 323–337. MR1629591
Klimko, L. A. and Nelson, P. I. (1978). On conditional least squares estimation for stochastic processes. *The Annals of Statistics* **6**, 629–642. MR0494770
Latour, A. (1997). The multivariate GINAR(p) process. *Advances in Applied Probability* **29**, 228–248. MR1432938
Latour, A. (1998). Existence and stochastic structure of a non-negative integer-valued autoregressive process. *Journal of Time Series Analysis* **19**, 439–455. MR1652193
Li, C., Wang, D. and Zhang, H. (2015). First-order mixed integer-valued autoregressive processes with zero-inflated generalized power series innovations. *Journal of the Korean Statistical Society* **44**, 232–246. MR3342635
Li, Q., Lian, H. and Zhu, F. (2016). Robust closed-form estimators for the integer-valued GARCH(1, 1) model. *Computational Statistics & Data Analysis* **101**, 209–225. MR3504846
Nicholls, D. F. and Quinn, B. G. (1982). *Random Coefficient Autoregressive Models: An Introduction*, Vol. **11**. Springer Science and Business Media. MR0671255
Steutel, F. W. and Van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *Annals of Probability* **7**, 893–899. MR0542141
Zhang, H. and Wang, D. (2015). Inference for random coefficient INAR(1) process based on frequency domain analysis. *Communications in Statistics Simulation and Computation* **44**, 1078–1100. MR3264922
Zheng, H. and Basawa, I. V. (2008). First-order observation-driven integer-valued autoregressive processes. *Statistics and Probability Letters* **78**, 1–9. MR2381267
Zheng, H., Basawa, I. V. and Datta, S. (2006). Inference for pth-order random coefficient integer-valued autoregressive processes. *Journal of Time Series Analysis* **27**, 411–440. MR2328539

Zheng, H., Basawa, I. V. and Datta, S. (2007). First-order random coefficient integer-valued autoregressive processes. *Journal of Statistical Planning and Inference* **137**, 212–229. MR2292852
Zhu, R. and Joe, H. (2006). Modelling count data time series with Markov processes based on binomial thinning. *Journal of Time Series Analysis* **27**, 725–738. MR2291211
Zikun, W. (1965). *Stochastic Process Theory*. Beijing: Scientific Press.

School of Mathematics
Jilin University
Changchun 130012
PR China
E-mail: lxf17@mails.jlu.edu.cn
wangdh@jlu.edu.cn