doi：10.13413/j.cnki.jdxblxb.2024524

具有相依 Bernoulli 计数序列的 RCINAR(p)模型的极大似然估计

李　琪，刘秀芳
（太原理工大学 数学学院，太原 030024）

摘要： 用具有相依 Bernoulli 计数序列的 p 阶随机系数整数值自回归 RCINAR(p)模型解决计数变量存在相关性特征数据的分析问题，得到了该模型的统计性质、参数条件极大似然估计及其渐近正态性，并通过实际数据分析验证了模型的有效性，精准捕捉了数据关联与趋势. 结果表明，该模型的估计量随样本量递增收敛于真值.

关键词： 相依 Bernoulli 计数序列；随机系数；条件极大似然估计；渐近正态性
中图分类号： O212.1　　**文献标志码：** A　　**文章编号：** 1671-5489(2025)05-1325-12

Maximum Likelihood Estimation of RCINAR (p) Model with Dependent Bernoulli Counting Series

LI Qi, LIU Xiufang
(*College of Mathematics, Taiyuan University of Technology, Taiyuan* 030024, *China*)

Abstract: We used a p-order random coefficient integer autoregressive RCINAR (p) model with dependent Bernoulli counting series to solve the analysis problem of data with correlated characteristics of counting variables. We obtained the statistical properties of the model, the conditional maximum likelihood estimation of the parameters, and their asymptotic normality, and the effectiveness of the model was verified through the analysis of actual data, accurately capturing data correlations and trends. These results show that the estimators of the model converge to the true values as the sample size increases.

Keywords: dependent Bernoulli counting series; random coefficient; conditional maximum likelihood estimation; asymptotic normality

整数时间序列模型在金融、经济、气象学等领域应用广泛. 在构建整数时间序列模型时，选择合适的稀疏算子至关重要. 文献[1]提出了二项稀疏算子“$\circ$”，对推动整数自回归模型的发展发挥了重要作用. 在此基础上，文献[2]提出了一阶整数时间序列自回归模型，文献[3]提出了负二项几何一阶整数自回归 NGINAR(1)过程，该过程使用了负二项稀疏算子. 由于单一稀疏算子模型无法准确表示实际数据，因此文献[4]提出了一种包含混合稀疏算子的整数自回归模型. 但上述模型中的稀疏系数假设恒定简化了研究. 而这些稀疏系数可能随时间变化，并可能受各种环境因素的影响，使其成

收稿日期： 2024-12-19.　**网络首发日期：** 2025-06-25.
第一作者简介： 李　琪(2001—)，女，汉族，硕士研究生，从事时间序列分析的研究，E-mail：1740538328@qq.com. **通信作者简介：** 刘秀芳(1990—)，女，汉族，博士，讲师，从事时间序列分析的研究，E-mail：liuxiufang@tyut.edu.cn.
基金项目： 山西省研究生创新创业项目(批准号：RC2400005695).
网络首发地址： https://link.cnki.net/urlid/22.1340.O.20250625.0949.001.

为随机变量，从而探索随机系数整数值自回归模型至关重要. 因此，文献[5]提出了一阶随机系数整数自回归过程；文献[6]深入研究了一种更高级的依赖驱动随机系数整数自回归模型 DDRCINAR(p).

随着模型系数随机化研究的发展，越来越多的整数自回归时间序列模型应运而生. 文献[7]对高阶依赖驱动随机系数整数值自回归时间序列模型 PoDDRCINAR(p)的严格平稳性、遍历性和参数估计方法进行了全面研究. 文献[8]研究了基于二项和负二项稀疏算子的混合模型 SDMINAR(p)的最大似然估计方法，并将其应用于 COVID-19 数据的分析. 文献[9]通过使用条件最小二乘法和条件加权最小二乘法对同一模型进行参数估计，扩展了上述研究结果. 上述研究都基于二项稀疏算子，展示了该模型在分析具有不同特征数据方面的多功能性. 文献[10]通过提出一阶随机系数混合稀疏算子整值自回归时间序列模型 RCMTINAR(1)，进一步推动了该领域研究的发展，该模型专门用于处理涉及可变特征元素计数的数据，利用混合稀疏算子进行有效的处理和分析. 文献[11]提出了高阶混合依赖驱动的随机系数整数值自回归时间序列模型 MDDRCINAR(p)，该模型基于二项和负二项稀疏算子. 上述研究对整数时间序列模型的发展和改进做出了贡献.

实际应用表明，一个事件的实现直接影响其他一些事件的实现或不实现. 因此，有必要建立包含相关计数序列的模型. 文献[12]引入了基于相关 Bernoulli 计数变量的广义稀疏算子，计数变量之间的相关性由特定参数确定，使该模型更易于应用. 在此基础上，本文提出具有相依 Bernoulli 计数序列的随机系数整数值自回归 RCINAR(p)模型. 首先，建立具有相依 Bernoulli 计数序列的随机系数整数值自回归 RCINAR(p)模型；其次，给出该模型参数的条件极大似然(conditional maximum likelihood, CML)估计方法，并研究其渐近性质；再次，通过模拟生成随机数据，深入研究这些估计量的性能及检验统计量；最后，用犯罪数据，引入 p 阶依赖驱动的随机系数整值自回归模型 DDRCINAR(p)模型与 RCINAR(p)模型对比，验证 RCINAR(p)模型的有效性和实用性.

1 模型结构设计

具有相依 Bernoulli 计数序列的随机系数整数值自回归 RCINAR(p)模型构造如下：

$$X_t=\sum_{i=1}^{p}\varphi_{ti}\odot_{\theta_i}X_{t-i}+\varepsilon_t,\qquad t\geqslant 1, \tag{1}$$

其中 $\varphi_{t1},\varphi_{t2},\cdots,\varphi_{tp}$ 的联合概率分布为

$$\begin{cases} P(\varphi_{t1}=\varphi_1,\varphi_{t2}=0,\cdots,\varphi_{tp}=0)=\alpha_1, \\ P(\varphi_{t1}=0,\varphi_{t2}=\varphi_2,\cdots,\varphi_{tp}=0)=\alpha_2, \\ \quad\vdots \\ P(\varphi_{t1}=0,\varphi_{t2}=0,\cdots,\varphi_{tp}=\varphi_p)=\alpha_p, \\ P(\varphi_{t1}=0,\varphi_{t2}=0,\cdots,\varphi_{tp}=0)=\alpha_0, \end{cases} \tag{2}$$

式中：$\varphi_i\in(0,1)$，$\alpha_0,\alpha_1,\cdots,\alpha_p$ 非负，且 $\sum\limits_{i=0}^{p}\alpha_i=1$；在给定 φ_{ti} 的条件下，

$$\varphi_{ti}\odot_{\theta_i}X_{t-i}\mid\varphi_{ti}=\sum_{l=1}^{X_{t-i}}V_{lt},\qquad V_{lt}=(1-V_l)W_{lt}+V_l\xi_t,$$

$\{V_l\}$是一个独立同分布的 Bernoulli 随机变量序列 Bernoulli(θ_i)，$\theta_i\in(0,1)$；在给定 φ_{ti} 下，$\{W_{lt}\}$和$\{\xi_t\}$分别是独立同分布的 Bernoulli 随机变量序列 Bernoulli(φ_{ti}). 根据文献[12]，$\varphi_{ti}\odot_{\theta_i}X_{t-i}\mid\varphi_{ti}$ 的分布为

$$\varphi_{ti}\odot_{\theta_i}X_{t-i}\mid\varphi_{ti}=\sum_{l=1}^{X_{t-i}}V_{lt}\overset{d}{=\!=}\begin{cases}\mathrm{Bin}(X_{t-i},\varphi_{ti}(1-\theta_i)), & \text{w. p. } 1-\varphi_{ti}, \\ \mathrm{Bin}(X_{t-i},\varphi_{ti}+\theta_i-\varphi_{ti}\theta_i)), & \text{w. p. } \varphi_{ti}.\end{cases}$$

假设序列$\{V_l\}$,$\{W_{lt}\}$和$\{\xi_t\}$对所有的 l 和 t 都独立. 此外，$\{\varepsilon_t\}$是独立同分布的非负整数值序列，并服从参数为 λ 的 Poisson 分布，且$\{\varphi_{ti},\ 1\leqslant i\leqslant p\}$和$\{\varepsilon_t\}$相互独立. 该模型参数为 $\boldsymbol{\eta}=(\varphi_1,\cdots,\varphi_p,\alpha_1,\cdots,\alpha_p,\theta_1,\cdots,\theta_p,\lambda)'$. 根据方程(2)，可得 φ_{ti} 的均值、方差及 φ_{ti} 与 φ_{tj}($i\neq j$)的协方差分别为

$$E(\varphi_{ti})=\alpha_i\varphi_i,\quad \mathrm{Var}(\varphi_{ti})=\alpha_i(1-\alpha_i)\varphi_i^2,\quad \mathrm{Cov}(\varphi_{ti},\varphi_{tj})=-\alpha_i\alpha_j\varphi_i\varphi_j.$$

定义 1　若 $\boldsymbol{X}=(X_1,X_2,\cdots,X_k)'_{k\times1}$，$\boldsymbol{X}_t=(X_t,X_{t-1},\cdots,X_{t-k+1})'_{k\times1}$，$\boldsymbol{X}_{t-1}=(X_{t-1},X_{t-2},\cdots,X_{t-k})'_{k\times1}$，$\boldsymbol{\varepsilon}_t=(\varepsilon_t,0,\cdots,0)'_{k\times1}$，

$$\boldsymbol{A}_t=\begin{pmatrix}\varphi_{t1} & \varphi_{t2} & \cdots & \varphi_{t,k-1} & \varphi_{tk}\\ 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & 0\end{pmatrix},\quad \boldsymbol{A}=\begin{pmatrix}\alpha_1\varphi_1 & \alpha_2\varphi_2 & \cdots & \alpha_{k-1}\varphi_{k-1} & \alpha_k\varphi_k\\ 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & 0\end{pmatrix},$$

$\boldsymbol{A}_t\odot_\theta\boldsymbol{X}=\left(\sum_{i=1}^{k}\varphi_{ti}\odot_{\theta_i}X_i,X_1,\cdots,X_{k-1}\right)'$，则 $\boldsymbol{X}_t=\boldsymbol{A}_t\odot_\theta\boldsymbol{X}_{t-1}+\boldsymbol{\varepsilon}_t$. 并且对于 $t\geqslant1$，$E(\boldsymbol{A}_t\odot_\theta\boldsymbol{X}\mid\boldsymbol{X})=\boldsymbol{A}\boldsymbol{X}$，$E(\boldsymbol{A}_t\odot_\theta\boldsymbol{X})=\boldsymbol{A}E(\boldsymbol{X})$.

命题 1　假设$\{X_t\}$是由方程(1)和(2)定义的 RCINAR(p)模型生成的序列，则其统计性质如下：

1) $E(X_t\mid X_{t-1},\cdots,X_{t-p})=\sum_{i=1}^{p}\alpha_i\varphi_iX_{t-i}+\lambda$；

2) 如果该模型是 k 阶平稳的，则 $E(X_t)=\lambda\left(1-\sum_{i=1}^{p}\alpha_i\varphi_i\right)^{-1}$；

3) $\mathrm{Var}(X_t\mid X_{t-1},\cdots,X_{t-p})=\sum_{i=1}^{p}\alpha_i\varphi_iX_{t-i}(1-\varphi_i)(1-\theta_i^2)+\sum_{i=1}^{p}\alpha_i\varphi_iX_{t-i}^2(\varphi_i+\theta_i^2-\varphi_i\theta_i^2)+\lambda-\left(\sum_{i=1}^{p}\alpha_i\varphi_iX_{t-i}\right)^2$；

4) 令 $\gamma_k=\mathrm{Cov}(X_t,X_{t-k})$，$\rho_k=\gamma_k/\gamma_0$，则 $\gamma_k=\sum_{i=1}^{p}\alpha_i\varphi_i\gamma_{k-i}$，$\rho_k=\sum_{i=1}^{p}\alpha_i\varphi_i\rho_{|k-i|}$.

证明：令 $S_i=\varphi_{ti}\odot_{\theta_i}X_{t-i}(i=1,2,\cdots,p)$，则 S_i 的条件矩母函数为

$$M_{S_i\mid\varphi_{ti},X_{t-i}}=(1-\varphi_{ti})[\varphi_{ti}(1-\theta_i)\mathrm{e}^t+1-\varphi_{ti}(1-\theta_i)]^{X_{t-i}}+\varphi_{ti}[(\varphi_{ti}+\theta_i-\varphi_{ti}\theta_i)\mathrm{e}^t+1-\varphi_{ti}-\theta_i+\varphi_{ti}\theta_i]^{X_{t-i}},$$

从而可得 S_i 的 k 阶条件矩为 $E(S_i^k\mid\varphi_{ti},X_{t-i})=M^{(k)}_{S_i\mid\varphi_{ti},X_{t-i}}(0)$. 因此

$$\begin{aligned}E(X_t^m\mid X_{t-1},\cdots,X_{t-p})=&\sum_{k=0}^{m-1}\mathrm{C}_m^k\sum_{i=1}^{p}E(E[(\varphi_{ti}\odot_{\theta_i}X_{t-i})^{m-k}\mid\varphi_{ti},X_{t-i}]\mid X_{t-i})E(\varepsilon_t^k)+E(\varepsilon_t^m)=\\&\sum_{k=0}^{m-1}\mathrm{C}_m^k\sum_{i=1}^{p}E(E(S_i^{m-k}\mid\varphi_{ti},X_{t-i})\mid X_{t-i})E(\varepsilon_t^k)+E(\varepsilon_t^m)=\\&\sum_{k=0}^{m-1}\mathrm{C}_m^k\sum_{i=1}^{p}E(M^{(m-k)}_{S_i\mid\varphi_{ti},X_{t-i}}(0)\mid X_{t-i})E(\varepsilon_t^k)+E(\varepsilon_t^m),\end{aligned}$$

根据期望和方差的公式，可推出统计性质 1)～ 3).

根据定义 1 和协方差原理，可得

$$\begin{aligned}\boldsymbol{\gamma}_k=&E[(\boldsymbol{X}_t-E(\boldsymbol{X}_t))(\boldsymbol{X}_{t-k}-E(\boldsymbol{X}_{t-k}))]=E(\boldsymbol{X}_t\boldsymbol{X}_{t-k})-E(\boldsymbol{X}_t)E(\boldsymbol{X}_{t-k})=\\&E((\boldsymbol{A}_t\odot_\theta\boldsymbol{X}_{t-1}+\boldsymbol{\varepsilon}_t)\boldsymbol{X}_{t-k})-E(\boldsymbol{X}_t)E(\boldsymbol{X}_{t-k})=\\&E((\boldsymbol{A}_t\odot_\theta\boldsymbol{X}_{t-1})\boldsymbol{X}_{t-k})+E(\boldsymbol{\varepsilon}_t\boldsymbol{X}_{t-k})-E(\boldsymbol{X}_t)E(\boldsymbol{X}_{t-k})=\\&\boldsymbol{A}E(\boldsymbol{X}_{t-1}\boldsymbol{X}_{t-k})+E(\boldsymbol{\varepsilon}_t\boldsymbol{X}_{t-k})-E(\boldsymbol{X}_t)E(\boldsymbol{X}_{t-k})=\\&\boldsymbol{A}\boldsymbol{\gamma}_{k-1}+\boldsymbol{A}E(\boldsymbol{X}_{t-1})E(\boldsymbol{X}_{t-k})+E(\boldsymbol{\varepsilon}_t\boldsymbol{X}_{t-k})-E(\boldsymbol{X}_t)E(\boldsymbol{X}_{t-k})=\\&\boldsymbol{A}\boldsymbol{\gamma}_{k-1}+E(\boldsymbol{A}_t\odot_\theta\boldsymbol{X}_{t-1}+\boldsymbol{\varepsilon}_t)E(\boldsymbol{X}_{t-k})-E(\boldsymbol{X}_t)E(\boldsymbol{X}_{t-k})=\boldsymbol{A}\boldsymbol{\gamma}_{k-1},\end{aligned}$$

所以 $\gamma_k=\sum_{i=1}^{p}\alpha_i\varphi_i\gamma_{k-i}$. 类似地，可得 $\boldsymbol{\rho}_k=\boldsymbol{A}\boldsymbol{\rho}_{k-1}$，$\rho_k=\sum_{i=1}^{p}\alpha_i\varphi_i\rho_{|k-i|}$. 证毕.

命题 2　由 RCINAR(p)模型生成的$\{X_t\}$序列是严平稳遍历的.

证明：根据文献[13]中定理 4.1 的证明或文献[14]的遍历性证明可知该命题成立.

2　参数推断方法

为得到参数 $\boldsymbol{\eta}$ 的条件极大似然估计量，给出 RCINAR(p)模型的转移概率为

$$P_{\boldsymbol{\eta}}(X_t \mid X_{t-1}, X_{t-2}, \cdots, X_{t-p}) = \sum_{i=1}^{p} \sum_{m_i=0}^{\min\{X_t, X_{t-i}\}} \alpha_i q(X_t, m_i)[(1-\varphi_i)Z(X_{t-i}, m_i) + \varphi_i Y(X_{t-i}, m_i)] + \left(1 - \sum_{i=1}^{p} \alpha_i\right) q(X_t, m_0), \tag{3}$$

其中

$$q(X_t, m_i) = \frac{\lambda^{X_t - m_i} \mathrm{e}^{-\lambda}}{(X_t - m_i)!}, \quad i = 0, 1, \cdots, p, \quad m_0 = 0, \tag{4}$$

$$Z(X_{t-i}, m_i) = \binom{X_{t-i}}{m_i} [\varphi_i(1-\theta_i)]^{m_i} [1 - \varphi_i(1-\theta_i)]^{X_{t-i} - m_i}, \tag{5}$$

$$Y(X_{t-i}, m_i) = \binom{X_{t-i}}{m_i} (\varphi_i + \theta_i - \varphi_i\theta_i)^{m_i} (1 - \varphi_i - \theta_i + \varphi_i\theta_i)^{X_{t-i} - m_i}. \tag{6}$$

假设 $X_1, X_2, \cdots, X_n$ 是由 RCINAR(p)过程生成的随机变量，其真实参数值为 $\boldsymbol{\eta}_0$，对数似然函数为 $l(\boldsymbol{\eta}) = \sum_{t=p+1}^{n} \log P_{\boldsymbol{\eta}}(X_t \mid X_{t-1}, X_{t-2}, \cdots, X_{t-p})$，其中 $P_{\boldsymbol{\eta}}(X_t \mid X_{t-1}, X_{t-2}, \cdots, X_{t-p})$由方程(3)给出，并简记为 $P_{\boldsymbol{\eta}}$. 则 RCINAR(p)模型的得分函数为

$$\frac{\partial l(\boldsymbol{\eta})}{\partial \alpha_i} = \sum_{t=p+1}^{n} \frac{1}{P_{\boldsymbol{\eta}}} \left\{ \sum_{m_i=0}^{\min\{X_t, X_{t-i}\}} q(X_t, m_i)[(1-\varphi_i)Z(X_{t-i}, m_i) + \varphi_i Y(X_{t-i}, m_i)] - q(X_t, m_0) \right\}, \tag{7}$$

$$\begin{aligned}\frac{\partial l(\boldsymbol{\eta})}{\partial \varphi_i} = \sum_{t=p+1}^{n} \frac{1}{P_{\boldsymbol{\eta}}} \Bigg\{ &\sum_{m_i=0}^{\min\{X_t, X_{t-i}-1\}} \alpha_i q(X_t, m_i)\big(Y(X_{t-i}, m_i) - Z(X_{t-i}, m_i) + \\ &(1-\varphi_i)(1-\theta_i)X_{t-i}[Z(X_{t-i}-1, m_i-1) - Z(X_{t-i}-1, m_i)] + \\ &\varphi_i(1-\theta_i)X_{t-i}[Y(X_{t-i}-1, m_i-1) - Y(X_{t-i}-1, m_i)]\big) \Bigg\},\end{aligned} \tag{8}$$

$$\begin{aligned}\frac{\partial l(\boldsymbol{\eta})}{\partial \theta_i} = \sum_{t=p+1}^{n} \frac{1}{P_{\boldsymbol{\eta}}} \Bigg\{ &\sum_{m_i=0}^{\min\{X_t, X_{t-i}-1\}} \alpha_i q(X_t, m_i)\varphi_i(1-\varphi_i)X_{t-i}[Z(X_{t-i}-1, m_i) - Z(X_{t-i}-1, m_i-1) + \\ &Y(X_{t-i}-1, m_i-1) - Y(X_{t-i}-1, m_i)] \Bigg\},\end{aligned} \tag{9}$$

$$\begin{aligned}\frac{\partial l(\boldsymbol{\eta})}{\partial \lambda} = \sum_{t=p+1}^{n} \frac{1}{P_{\boldsymbol{\eta}}} \Bigg\{ &\sum_{i=1}^{p} \sum_{m_i=0}^{\min\{X_t-1, X_{t-i}\}} \alpha_i q(X_t-1, m_i)[(1-\varphi_i)Z(X_{t-i}, m_i) + \varphi_i Y(X_{t-i}, m_i)] + \\ &(1 - \sum_{i=1}^{p} \alpha_i) q(X_t-1, m_0) - P_{\boldsymbol{\eta}} \Bigg\}.\end{aligned} \tag{10}$$

所以，可求解令得分函数为零的方程组推出未知参数 $\boldsymbol{\eta}$ 的条件极大似然估计量 $\hat{\boldsymbol{\eta}}$.

定义 2 若 $R(X_{t-1}, k_1) = \binom{X_{t-1}}{k_1} \beta^{k_1}(1-\beta)^{X_{t-1}-k_1}$，则

$$R(X_{t-1}, k_1) = \beta R(X_{t-1}-1, k_1-1) + (1-\beta)R(X_{t-1}-1, k_1),$$

$$\frac{\partial R(X_{t-1}, k_1)}{\partial \beta} = \frac{X_{t-1}}{1-\beta}[R(X_{t-1}-1, k_1-1) - R(X_{t-1}, k_1)].$$

通过定义 2，可给出 $Z(X_{t-i}, m_i)$和 $Y(X_{t-i}, m_i)$对 φ_i 偏导的关系式如下：

$$\frac{\partial Z(X_{t-i}, m_i)}{\partial \varphi_i} = \frac{X_{t-i}(1-\theta_i)}{1-\varphi_i(1-\theta_i)}[Z(X_{t-i}-1, m_i-1) - Z(X_{t-i}, m_i)], \tag{11}$$

$$\frac{\partial Y(X_{t-i}, m_i)}{\partial \varphi_i} = \frac{X_{t-i}(1-\theta_i)}{1-\varphi_i-\theta_i+\varphi_i\theta_i}[Y(X_{t-i}-1, m_i-1) - Y(X_{t-i}, m_i)], \tag{12}$$

$$\frac{\partial \log P_{\boldsymbol{\eta}}}{\partial \varphi_i} = \frac{1}{P_{\boldsymbol{\eta}}} \sum_{m_i=0}^{\min\{X_t, X_{t-i}-1\}} \alpha_i q(X_t, m_i)\Big[Y(X_{t-i}, m_i) - Z(X_{t-i}, m_i) + (1-\varphi_i)\frac{\partial Z(X_{t-i}, m_i)}{\partial \varphi_i} + \varphi_i \frac{\partial Y(X_{t-i}, m_i)}{\partial \varphi_i}\Big], \tag{13}$$

用于证明参数估计量的渐近正态性.

命题 3

$$-\max\left\{\frac{X_{t-i}(1-\theta_i)}{1-\varphi_i(1-\theta_i)}+\frac{1}{1-\varphi_i},\frac{X_{t-i}(1-\theta_i)}{1-\varphi_i-\theta_i+\varphi_i\theta_i}-\frac{1}{\varphi_i}\right\}<\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\varphi_i}<$$
$$\max\left\{\frac{X_{t-i}}{\varphi_i-\varphi_i^2(1-\theta_i)},\frac{X_{t-i}(1-\theta_i)}{(1-\varphi_i-\theta_i+\varphi_i\theta_i)(\varphi_i+\theta_i-\varphi_i\theta_i)}\right\}. \tag{14}$$

证明：令

$$G=\frac{X_{t-i}(1-\varphi_i)(1-\theta_i)}{1-\varphi_i(1-\theta_i)},\quad H=\frac{X_{t-i}(1-\varphi_i)(1-\theta_i)}{1-\varphi_i-\theta_i+\varphi_i\theta_i},\quad \vartheta_1=\varphi_i(1-\theta_i),\quad \vartheta_2=\varphi_i+\theta_i-\varphi_i\theta_i,$$

则

$$\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\varphi_i}=\frac{1}{P_{\boldsymbol{\eta}}}\sum_{m_i=0}^{\min\{X_t,X_{t-i}-1\}}\alpha_i q(X_t,m_i)\{Y(X_{t-i},m_i)-Z(X_{t-i},m_i)+$$
$$G[Z(X_{t-i}-1,m_i-1)-Z(X_{t-i},m_i)]+H[Y(X_{t-i}-1,m_i-1)-Y(X_{t-i},m_i)]\}. \tag{15}$$

首先，证明不等式(14)左侧成立，

$$\text{式(15)}=\frac{1}{P_{\boldsymbol{\eta}}}\sum_{m_i=0}^{\min\{X_t,X_{t-i}-1\}}\alpha_i q(X_t,m_i)[GZ(X_{t-i}-1,m_i-1)+HY(X_{t-i}-1,m_i-1)-$$
$$(G+1)Z(X_{t-i},m_i)-(H-1)Y(X_{t-i},m_i)]>$$
$$\frac{1}{P_{\boldsymbol{\eta}}}\sum_{m_i=0}^{\min\{X_t,X_{t-i}-1\}}\alpha_i q(X_t,m_i)[-(G+1)Z(X_{t-i},m_i)-(H-1)Y(X_{t-i},m_i)]>$$
$$-\max\left\{\frac{G}{1-\varphi_i}+\frac{1}{1-\varphi_i},\frac{H}{\varphi_i}-\frac{H-1}{\varphi_i}\right\}.$$

其次，证明不等式(14)右侧成立，

$$\text{式(15)}=\frac{1}{P_{\boldsymbol{\eta}}}\sum_{m_i=0}^{\min\{X_t,X_{t-i}-1\}}\alpha_i q(X_t,m_i)[GZ(X_{t-i}-1,m_i-1)+HY(X_{t-i}-1,m_i-1)-$$
$$(G+1)Z(X_{t-i},m_i)-(H-1)Y(X_{t-i},m_i)]<$$
$$\frac{1}{P_{\boldsymbol{\eta}}}\sum_{m_i=0}^{\min\{X_t,X_{t-i}-1\}}\alpha_i q(X_t,m_i)[GZ(X_{t-i}-1,m_i-1)+HY(X_{t-i}-1,m_i-1)]=$$
$$\frac{1}{P_{\boldsymbol{\eta}}}\sum_{m_i=0}^{\min\{X_t,X_{t-i}-1\}}\alpha_i q(X_t,m_i)\left\{\frac{G}{\vartheta_1}[Z(X_{t-i},m_i)-(1-\vartheta_1)Z(X_{t-i}-1,m_i)]+\right.$$
$$\left.\frac{H}{\vartheta_2}[Y(X_{t-i},m_i)-(1-\vartheta_2)Y(X_{t-i}-1,m_i)]\right\}<$$
$$\frac{1}{P_{\boldsymbol{\eta}}}\sum_{m_i=0}^{\min\{X_t,X_{t-i}-1\}}\alpha_i q(X_t,m_i)\left\{\frac{G}{\vartheta_1}Z(X_{t-i},m_i)+\frac{H}{\vartheta_2}Y(X_{t-i},m_i)\right\}<$$
$$\max\left\{\frac{G}{\vartheta_1(1-\varphi_i)},\frac{H}{\vartheta_2\varphi_i}\right\}.$$

证毕.

下面证明参数估计量 $\boldsymbol{\eta}$ 的渐近正态性，首先给出如下正则条件[15]：

(i) 集合$\left\{m: P(\varepsilon_t=m)=f(m,\lambda)=\dfrac{\lambda^m e^{-\lambda}}{m!}>0\right\}$不依赖于 λ，确保概率计算过程的一致性和稳定性；

(ii) $E(\varepsilon_t^3)=\lambda^3+3\lambda^2+\lambda<\infty$，提供有界性条件，避免计算过程出现无穷大的情况；

(iii) $P(\varepsilon_t=m)$是三次连续可微的；

(iv) 对任意的 $\lambda'\in B$(这里 B 是集合 $\mathbb{R}$ 的一个开子集)，存在 λ'的一个领域 U，使得

$$\sum_{m=0}^{\infty}\sup_{\lambda\in U}f(m,\lambda)<\infty,\quad \sum_{m=0}^{\infty}\sup_{\lambda\in U}\left|\frac{\partial f(m,\lambda)}{\partial\lambda}\right|<\infty,\quad \sum_{m=0}^{\infty}\sup_{\lambda\in U}\left|\frac{\partial^2 f(m,\lambda)}{\partial\lambda^2}\right|<\infty;$$

(v) 对于$\lambda'\in B$，存在λ'的一个领域U和常数C_1,C_2,C_3，使得序列$\phi_1(n)=C_1n$，$\phi_2(n)=C_2n^2$，$\phi_3(n)=C_3n^3$，当$n\geqslant 0$时，对任意的$\lambda\in U$和$m\leqslant n$，在过程$\{X_t\}$具有平稳遍历分布的条件下，ε_t的非零概率密度函数$f(m,\lambda)$满足

$$\left|\frac{\partial f(m,\lambda)}{\partial\lambda}\right|<\phi_1(n)f(m,\lambda),\quad \left|\frac{\partial^2 f(m,\lambda)}{\partial\lambda^2}\right|<\phi_2(n)f(m,\lambda),\quad \left|\frac{\partial^3 f(m,\lambda)}{\partial\lambda^3}\right|<\phi_3(n)f(m,\lambda);$$

(vi) Fisher信息矩阵$\boldsymbol{I}(\boldsymbol{\eta})=(\sigma_{ij})_{(3k+1)\times(3k+1)}(i,j=1,2,\cdots,3k+1)$，且$\boldsymbol{I}(\boldsymbol{\eta})$是非奇异的，并有$\boldsymbol{I}(\boldsymbol{\eta})=E\left(\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\boldsymbol{\eta}_i}\cdot\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\boldsymbol{\eta}_j}\right)$.

定理1 假设正则条件(i)～(vi)成立，则$\boldsymbol{\eta}$的条件极大似然估计量$\hat{\boldsymbol{\eta}}$具有渐近正态性，且有$\sqrt{n}(\hat{\boldsymbol{\eta}}-\boldsymbol{\eta})\xrightarrow{\text{d}}N(\boldsymbol{0},\boldsymbol{I}^{-1}(\boldsymbol{\eta}))$，$n\to\infty$.

证明：根据文献[16]，显然满足条件(i)～(iv). 又随机项ε_t服从参数为λ的Poisson分布，因此也满足条件(v). 所以只需证明满足条件(vi)即可，从而只需证明下列结论成立：

1) $E\left|\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\varphi_i}\right|^2<\infty$;

2) $E\left|\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\theta_i}\right|^2<\infty$;

3) $E\left|\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\alpha_i}\right|^2<\infty$;

4) $E\left|\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\lambda}\right|^2<\infty$;

5) $E\left|\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\varphi_i}\cdot\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\varphi_j}\right|<\infty$，$j\neq i$，$i,j=1,2,\cdots,p$;

6) $E\left|\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\varphi_i}\cdot\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\theta_j}\right|<\infty$;

7) $E\left|\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\varphi_i}\cdot\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\alpha_j}\right|<\infty$;

8) $E\left|\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\varphi_i}\cdot\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\lambda}\right|<\infty$;

9) $E\left|\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\theta_i}\cdot\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\theta_j}\right|<\infty$;

10) $E\left|\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\theta_i}\cdot\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\alpha_j}\right|<\infty$;

11) $E\left|\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\theta_i}\cdot\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\lambda}\right|<\infty$;

12) $E\left|\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\alpha_i}\cdot\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\alpha_j}\right|<\infty$;

13) $E\left|\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\alpha_i}\cdot\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\lambda}\right|<\infty$.

先证明结论1)，根据得分函数(8)和命题3，可得

$$\left|\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\varphi_i}\right|\leqslant\max\left\{\frac{X_{t-i}(1-\theta_i)}{1-\varphi_i(1-\theta_i)}+\frac{1}{1-\varphi_i},\ \frac{X_{t-i}(1-\theta_i)}{1-\varphi_i-\theta_i+\varphi_i\theta_i}-\frac{1}{\varphi_i},\right.$$
$$\left.\frac{X_{t-i}}{\varphi_i-\varphi_i^2(1-\theta_i)},\frac{X_{t-i}(1-\theta_i)}{(1-\varphi_i-\theta_i+\varphi_i\theta_i)(\varphi_i+\theta_i-\varphi_i\theta_i)}\right\}<C\cdot X_{t-i},$$

则$E\left|\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\varphi_i}\right|^2<C\cdot E(X_{t-i}^2)<\infty$. 结论2)同理结论1)，可得$\left|\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\theta_i}\right|<C\cdot X_{t-i}<\infty$，则$E\left|\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\theta_i}\right|^2<\infty$. 结论3)根据得分函数(7)，可得$E\left|\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\alpha_i}\right|<\frac{1}{\alpha_i}<\infty$，从而$E\left|\frac{\partial\log P_{\boldsymbol{\eta}}}{\partial\alpha_i}\right|^2<\infty$. 结论4)根据条件(v)，有

$$\left|\frac{\partial \log P_{\boldsymbol{\eta}}}{\partial \lambda}\right| = \left|\frac{1}{P_{\boldsymbol{\eta}}} \cdot \frac{\partial P_{\boldsymbol{\eta}}}{\partial \lambda}\right| < \left|\frac{q(X_t - 1, m_i)}{q(X_t, m_i)}\right| < \frac{X_t}{\lambda} = \phi_1(X_t),$$

则 $E\left|\frac{\partial \log P_{\boldsymbol{\eta}}}{\partial \lambda}\right|^2 < E\phi_1^2(X_t) < \infty$. 最后，基于结论 1)～ 4)和条件(v)，可得结论 5)～13)成立，从而可定义 Fisher 信息矩阵 $\boldsymbol{I}(\boldsymbol{\eta})$，同时根据条件(vi)，可得 $\boldsymbol{I}(\boldsymbol{\eta})$是非奇异的，证毕.

3　仿真实验与验证

下面通过仿真实验评估 RCINAR(2)模型条件极大似然估计量的样本性能. 采用估计量的有限样本均值(Mean)、平均绝对偏差误差(MADE)、均方误差(MSE)和标准差(SD)4 个指标进行评价，其中

$$\mathrm{MADE} = \frac{1}{M}\sum_{i=1}^{M} |\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}|,$$

$$\mathrm{MSE} = \frac{1}{M}\sum_{i=1}^{M} (\hat{\boldsymbol{\eta}} - \boldsymbol{\eta})^2,$$

$$\mathrm{SD} = \sqrt{\frac{1}{M-1}\sum_{i=1}^{M} (\hat{\boldsymbol{\eta}} - \bar{\boldsymbol{\eta}})^2},$$

这里 $\hat{\boldsymbol{\eta}}$ 表示 $\boldsymbol{\eta}$ 的估计量，$\bar{\boldsymbol{\eta}}$ 表示 $\hat{\boldsymbol{\eta}}$ 的平均值，M 表示重复模拟次数. 考虑下列 4 组参数组合：

第一组参数：$\alpha_1 = 0.3$，$\alpha_2 = 0.4$，$\varphi_1 = 0.6$，$\varphi_2 = 0.3$，$\theta_1 = 0.2$，$\theta_2 = 0.4$，$\lambda = 1$；

第二组参数：$\alpha_1 = 0.3$，$\alpha_2 = 0.4$，$\varphi_1 = 0.6$，$\varphi_2 = 0.3$，$\theta_1 = 0.2$，$\theta_2 = 0.4$，$\lambda = 0.7$；

第三组参数：$\alpha_1 = 0.3$，$\alpha_2 = 0.3$，$\varphi_1 = 0.4$，$\varphi_2 = 0.4$，$\theta_1 = 0.2$，$\theta_2 = 0.4$，$\lambda = 1$；

第四组参数：$\alpha_1 = 0.3$，$\alpha_2 = 0.3$，$\varphi_1 = 0.4$，$\varphi_2 = 0.4$，$\theta_1 = 0.4$，$\theta_2 = 0.4$，$\lambda = 0.7$.

样本量的取值分别为 100，300，500，800，1 000，实验重复模拟 500 次. 图 1 为样本量为 100 时，RCINAR(2)模型分别在 4 组参数下的路径图.

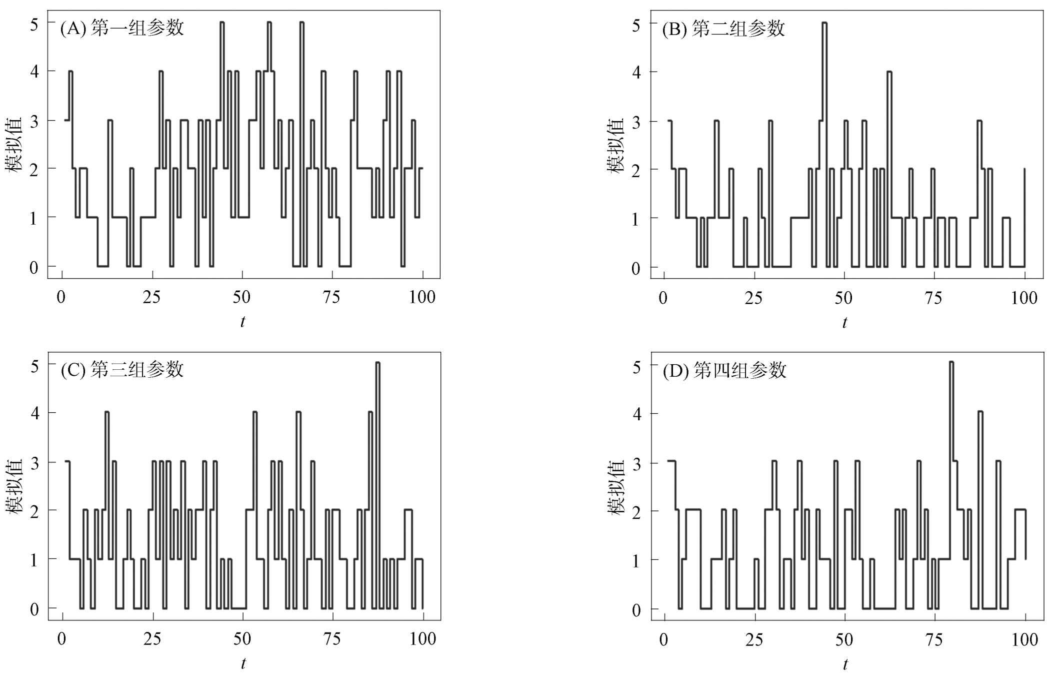

图 1　样本量为 100 时，RCINAR(2)模型在 4 组参数下的路径图

Fig. 1　Path diagrams of RCINAR(2) model under four sets of parameters with sample size of 100

表 1～表 4 分别为 4 组参数组合在不同样本量下参数估计的结果. 由表 1～表 4 可见，随着样本量的增加，Mean 逐渐趋于真值；MADE 和 MSE 逐渐趋于 0，表明了条件极大似然估计方法的稳定性；SD 逐渐趋于 0，模拟重复中得到的估计值相对集中，表明了条件极大似然估计方法的可靠性.

表1 第一组参数在不同样本量下参数估计的结果

Table 1 Results for parameter estimation of the first set of parameters at different sample sizes

n	评价指标	α_1	α_2	φ_1	φ_2	θ_1	θ_2	λ
100	Mean	0.299 9	0.387 0	0.508 7	0.328 9	0.633 9	0.567 9	1.017 0
	MADE	0.032 6	0.042 5	0.154 9	0.108 2	0.445 7	0.267 1	0.093 3
	MSE	0.002 7	0.004 4	0.040 6	0.018 3	0.243 5	0.096 4	0.013 9
	SD	0.051 9	0.065 0	0.179 7	0.132 4	0.256 9	0.261 4	0.117 0
300	Mean	0.294 9	0.398 9	0.552 5	0.329 2	0.560 2	0.490 8	1.001 1
	MADE	0.018 4	0.025 5	0.091 8	0.064 4	0.379 3	0.201 1	0.053 5
	MSE	0.000 6	0.001 3	0.014 4	0.006 5	0.188 0	0.059 1	0.004 4
	SD	0.024 8	0.035 6	0.110 1	0.075 0	0.241 7	0.225 8	0.066 2
500	Mean	0.294 3	0.400 2	0.559 4	0.324 6	0.485 5	0.464 2	1.000 0
	MADE	0.017 0	0.022 4	0.071 7	0.052 4	0.302 8	0.166 4	0.039 2
	MSE	0.000 5	0.000 9	0.008 7	0.004 1	0.124 6	0.042 3	0.002 5
	SD	0.021 5	0.029 6	0.084 1	0.058 9	0.207 8	0.195 6	0.050 2
800	Mean	0.295 4	0.400 2	0.571 5	0.316 6	0.438 6	0.431 2	1.001 8
	MADE	0.014 6	0.019 7	0.053 4	0.042 0	0.254 5	0.143 1	0.034 2
	MSE	0.000 4	0.000 6	0.004 8	0.002 8	0.092 1	0.030 5	0.001 9
	SD	0.018 2	0.025 1	0.063 3	0.050 2	0.187 7	0.171 9	0.043 2
1 000	Mean	0.297 6	0.399 8	0.580 3	0.312 4	0.436 5	0.464 4	1.000 3
	MADE	0.012 8	0.100 4	0.052 8	0.047 5	0.252 4	0.161 3	0.028 5
	MSE	0.000 3	0.000 5	0.004 4	0.002 1	0.085 8	0.028 1	0.001 3
	SD	0.016 2	0.024 7	0.063 3	0.044 7	0.172 9	0.140 7	0.036 2

表2 第二组参数在不同样本量下参数估计的结果

Table 2 Results for parameter estimation of the second set of parameters at different sample sizes

n	评价指标	α_1	α_2	φ_1	φ_2	θ_1	θ_2	λ
100	Mean	0.301 3	0.392 0	0.515 0	0.331 9	0.620 1	0.580 1	0.709 2
	MADE	0.028 4	0.035 6	0.158 3	0.113 0	0.434 5	0.275 0	0.078 7
	MSE	0.012 0	0.008 1	0.041 5	0.020 2	0.241 9	0.105 4	0.010 0
	SD	0.044 8	0.055 6	0.185 4	0.138 7	0.256 0	0.270 4	0.099 5
300	Mean	0.309 6	0.407 7	0.544 1	0.320 1	0.611 2	0.538 0	0.704 2
	MADE	0.027 7	0.030 4	0.102 8	0.078 4	0.422 6	0.236 0	0.044 8
	MSE	0.007 9	0.005 7	0.017 8	0.009 6	0.224 3	0.078 9	0.003 2
	SD	0.038 7	0.035 1	0.121 1	0.096 0	0.235 2	0.244 9	0.056 3
500	Mean	0.297 6	0.397 4	0.560 9	0.316 7	0.554 5	0.500 8	0.701 4
	MADE	0.012 6	0.017 8	0.080 5	0.059 7	0.366 7	0.193 7	0.036 8
	MSE	0.000 3	0.000 6	0.010 7	0.005 3	0.174 8	0.056 7	0.002 1
	SD	0.017 4	0.024 8	0.096 1	0.071 1	0.221 8	0.216 0	0.045 9
800	Mean	0.298 4	0.396 4	0.575 1	0.313 1	0.489 2	0.448 6	0.702 1
	MADE	0.012 2	0.016 4	0.060 4	0.045 8	0.313 3	0.163 0	0.028 0
	MSE	0.000 3	0.000 5	0.006 3	0.003 2	0.133 0	0.040 4	0.001 2
	SD	0.015 8	0.021 4	0.075 2	0.054 6	0.220 3	0.195 2	0.035 2
1 000	Mean	0.297 0	0.399 0	0.576 4	0.310 8	0.473 7	0.447 6	0.699 6
	MADE	0.011 6	0.015 1	0.058 5	0.042 5	0.283 4	0.161 1	0.024 5
	MSE	0.000 2	0.000 4	0.005 5	0.003 0	0.107 8	0.038 7	0.000 9
	SD	0.014 9	0.019 9	0.070 0	0.053 4	0.181 5	0.191 1	0.030 6

表 3　第三组参数在不同样本量下参数估计的结果

Table 3　Results for parameter estimation of the third set of parameters at different sample sizes

n	评价指标	α_1	α_2	φ_1	φ_2	θ_1	θ_2	λ
100	Mean	0.292 0	0.297 5	0.379 8	0.403 9	0.607 5	0.608 1	1.003 7
	MADE	0.035 7	0.031 5	0.139 1	0.137 3	0.430 0	0.281 5	0.088 0
	MSE	0.003 5	0.002 6	0.029 0	0.027 7	0.232 0	0.107 5	0.012 4
	SD	0.058 3	0.051 3	0.169 4	0.166 5	0.257 2	0.253 6	0.111 3
300	Mean	0.297 1	0.295 5	0.394 4	0.392 6	0.553 2	0.566 9	1.000 5
	MADE	0.020 2	0.020 6	0.090 8	0.091 0	0.367 8	0.230 4	0.051 0
	MSE	0.000 9	0.000 8	0.012 6	0.013 1	0.178 8	0.072 9	0.004 3
	SD	0.029 2	0.028 8	0.112 3	0.114 5	0.232 6	0.212 4	0.065 6
500	Mean	0.298 0	0.296 2	0.398 9	0.389 7	0.501 3	0.525 7	1.001 9
	MADE	0.017 6	0.019 0	0.072 6	0.076 6	0.322 7	0.202 5	0.038 8
	MSE	0.000 6	0.000 7	0.008 3	0.008 8	0.139 4	0.059 5	0.002 4
	SD	0.023 9	0.026 8	0.091 0	0.093 5	0.220 6	0.209 3	0.049 4
800	Mean	0.297 9	0.297 8	0.395 8	0.399 2	0.435 8	0.491 2	0.999 7
	MADE	0.015 5	0.014 9	0.061 0	0.058 7	0.255 7	0.174 4	0.031 4
	MSE	0.000 4	0.000 4	0.005 7	0.005 2	0.092 6	0.044 6	0.001 5
	SD	0.019 8	0.019 2	0.075 2	0.072 0	0.192 5	0.190 6	0.038 8
1 000	Mean	0.297 2	0.298 4	0.395 5	0.401 7	0.416 0	0.463 3	1.001 3
	MADE	0.014 8	0.014 2	0.052 7	0.051 8	0.239 8	0.157 4	0.027 6
	MSE	0.000 3	0.000 3	0.004 3	0.004 3	0.085 0	0.037 4	0.001 2
	SD	0.017 9	0.017 7	0.065 1	0.065 7	0.186 1	0.182 9	0.034 9

表 4　第四组参数在不同样本量下参数估计的结果

Table 4　Results for parameter estimation of the fourth set of parameters at different sample sizes

n	评价指标	α_1	α_2	φ_1	φ_2	θ_1	θ_2	λ
100	Mean	0.300 3	0.288 9	0.397 3	0.391 6	0.629 7	0.606 6	0.717 5
	MADE	0.031 9	0.031 0	0.153 7	0.148 1	0.281 2	0.278 1	0.078 7
	MSE	0.003 3	0.002 2	0.034 5	0.032 5	0.107 0	0.105 3	0.009 8
	SD	0.057 5	0.046 1	0.185 8	0.180 2	0.233 1	0.250 4	0.097 3
300	Mean	0.296 9	0.297 4	0.384 2	0.395 4	0.596 3	0.592 0	0.706 6
	MADE	0.016 8	0.017 0	0.105 3	0.106 3	0.258 4	0.246 2	0.040 1
	MSE	0.000 6	0.000 6	0.017 2	0.017 3	0.089 1	0.081 8	0.002 6
	SD	0.024 1	0.024 6	0.130 5	0.131 5	0.225 1	0.212 2	0.051 1
500	Mean	0.299 8	0.296 0	0.396 2	0.389 6	0.571 0	0.577 3	0.701 6
	MADE	0.013 1	0.014 4	0.080 3	0.079 4	0.230 0	0.228 8	0.034 2
	MSE	0.000 3	0.000 5	0.009 9	0.010 2	0.074 1	0.073 7	0.001 9
	SD	0.017 9	0.020 9	0.099 6	0.100 5	0.212 0	0.205 7	0.043 2
800	Mean	0.299 0	0.297 7	0.398 1	0.390 7	0.539 0	0.523 7	0.701 9
	MADE	0.010 9	0.011 2	0.064 4	0.065 9	0.199 4	0.200 3	0.027 2
	MSE	0.000 2	0.000 2	0.006 5	0.006 9	0.056 3	0.056 4	0.001 1
	SD	0.014 3	0.014 7	0.080 4	0.082 7	0.192 5	0.202 9	0.033 4
1000	Mean	0.299 0	0.298 1	0.396 6	0.395 9	0.517 1	0.514 1	0.701 9
	MADE	0.010 6	0.010 6	0.060 0	0.058 9	0.195 8	0.193 8	0.024 2
	MSE	0.000 2	0.000 2	0.005 5	0.005 4	0.053 8	0.053 7	0.000 9
	SD	0.013 7	0.013 7	0.073 9	0.073 8	0.180 3	0.201 9	0.029 9

图 2 为当样本量为 1 000 时 4 组参数组合得到的参数估计值 Q-Q 图[17]，其中横坐标表示基于理论正态分布计算出的分位数值，纵坐标表示模型参数估计值的分位数值. 由图 2 可见，所有参数的 CML 估计量都遵循渐近正态分布. 因此，本文结果与参数估计量的渐近正态性定理一致.

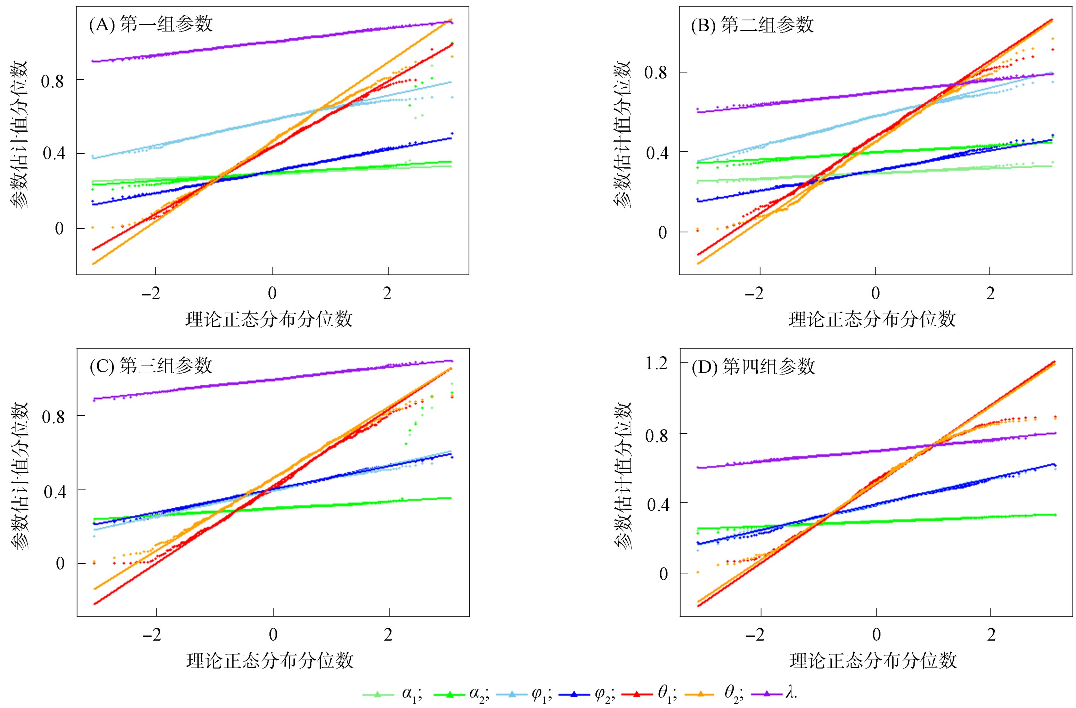

图 2　当样本量为 1 000 时不同参数组合参数估计值的 Q-Q 图

Fig. 2　Q-Q plots of parameter estimates for different parameter combinations when sample size is 1 000

4　实际数据对比研究

下面通过分析从 2013 年 3 月到 2024 年 9 月共 139 个在押的原住民群体中的成年女性犯罪数据，展示所构建 RCINAR(p)模型的实际应用效果. 该组数据来源于澳大利亚新南威尔士州犯罪统计与研究局(NSW Inverell of Crime Statistics and Research)发布的公开信息(网址为 https://www.bocsar.nsw.gov.au/). 该组数据的均值和方差分别为 2.323 7 和 4.713 3. 图 3 为该组数据的计数图、自相关(ACF)图和偏相关(PACF)图. 引入 p 阶依赖驱动的随机系数整值自回归模型 DDRCINAR(p)模型[6]与 RCINAR(p)模型进行对比，DDRCINAR(p)模型为

$$X_t = \sum_{i=1}^{p} \varphi_{ti} \circ X_{t-i} + \varepsilon_t, \qquad t \geqslant 1,$$

其中 $\varepsilon_t \sim \mathrm{Po}(\lambda)$，$\varphi_{ti} \circ X_{t-i} \,|\, \varphi_{ti} \sim \mathrm{Bin}(X_{t-i}, \varphi_{ti})$. 由图 3 的 PACF 图可知，模型应该考虑的阶数为 1，2，4 阶. 故本文考虑下列 6 种模型：RCINAR(1)，RCINAR(2)，RCINAR(4)，DDRCINAR(1)，DDRCINAR(2)和 DDRCINAR(4).

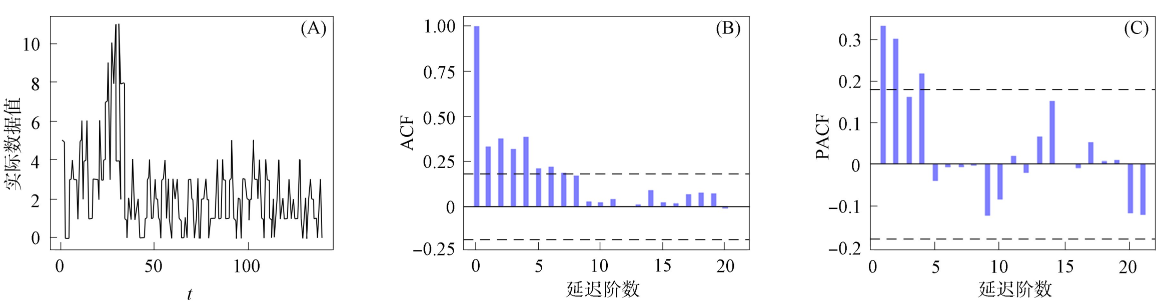

图 3　犯罪数据的计数图(A)、ACF 图(B)和 PACF 图(C)

Fig. 3　Count plot (A), ACF plot (B), and PACF plot (C) of crime data

表 5 列出了上述 6 种模型参数的条件极大似然估计值和用于评价模型优劣的 Akaike 信息准则(AIC)、Bayes 信息准则(BIC)、Hannan-Quinn 准则(HQ)和均方根误差准则(RMS). 由表 5 可见，

本文提出的 RCINAR(4)模型分析该组数据的 AIC，BIC，HQ 和 RMS 最小．基于 RCINAR(4)模型，图 4为拟合残差的路径图、ACF 图、PACF 图和累计周期图．由图 4 可见，该模型的拟合残差近似服从标准正态分布，表明本文提出的 RCINAR(4)模型更适合分析该组数据．

表 5　两种模型不同阶数对犯罪数据的参数极大似然估计和 AIC，BIC，HQ，RMS 值

Table 5　Maximum likelihood estimation of parameters and AIC, BIC, HQ, and RMS values of two models with different orders for crime data

RCINAR(1)	RCINAR(2)	RCINAR(4)	DDRCINAR(1)	DDRCINAR(2)	DDRCINAR(4)
$\hat{\alpha}_1=0.5032$	$\hat{\alpha}_1=0.3498$	$\hat{\alpha}_1=0.3740$	$\tilde{\alpha}_1=0.3343$	$\tilde{\alpha}_1=0.5859$	$\tilde{\alpha}_1=0.6077$
$\hat{\varphi}_1=0.4243$	$\hat{\alpha}_2=0.4322$	$\hat{\alpha}_2=0.1003$	$\tilde{\varphi}_1=0.5555$	$\tilde{\alpha}_2=0.1881$	$\tilde{\alpha}_2=0.0578$
$\hat{\theta}_1=0.7758$	$\hat{\varphi}_1=0.5487$	$\hat{\alpha}_3=0.2264$	$\tilde{\lambda}=1.7483$	$\tilde{\varphi}_1=0.3238$	$\tilde{\alpha}_3=0.1771$
$\hat{\lambda}=1.7226$	$\hat{\varphi}_2=0.2821$	$\hat{\alpha}_4=0.1621$		$\tilde{\varphi}_2=0.7243$	$\tilde{\alpha}_4=0.2043$
	$\hat{\theta}_1=0.6207$	$\hat{\varphi}_1=0.2252$		$\tilde{\lambda}=1.3211$	$\tilde{\varphi}_1=0.0901$
	$\hat{\theta}_2=0.6997$	$\hat{\varphi}_2=0.4007$			$\tilde{\varphi}_2=0.9817$
	$\hat{\lambda}=1.3912$	$\hat{\varphi}_3=0.4275$			$\tilde{\varphi}_3=0.0472$
		$\hat{\varphi}_4=0.7349$			$\tilde{\varphi}_4=0.7632$
		$\hat{\theta}_1=0.7685$			$\tilde{\lambda}=1.6181$
		$\hat{\theta}_2=0.9651$			
		$\hat{\theta}_3=0.7668$			
		$\hat{\theta}_4=0.2660$			
		$\hat{\lambda}=1.4515$			
AIC=542.9844	AIC=529.2562	AIC=525.4564	AIC=548.0525	AIC=531.3787	AIC=530.4796
BIC=554.6935	BIC=549.6961	BIC=543.2249	BIC=556.8342	BIC=545.9786	BIC=545.6271
HQ=541.3636	HQ=526.4093	HQ=520.1304	HQ=546.8368	HQ=529.3452	HQ=525.7924
RMS=2.0519	RMS=1.9945	RMS=1.9513	RMS=2.0625	RMS=1.9961	RMS=1.9771

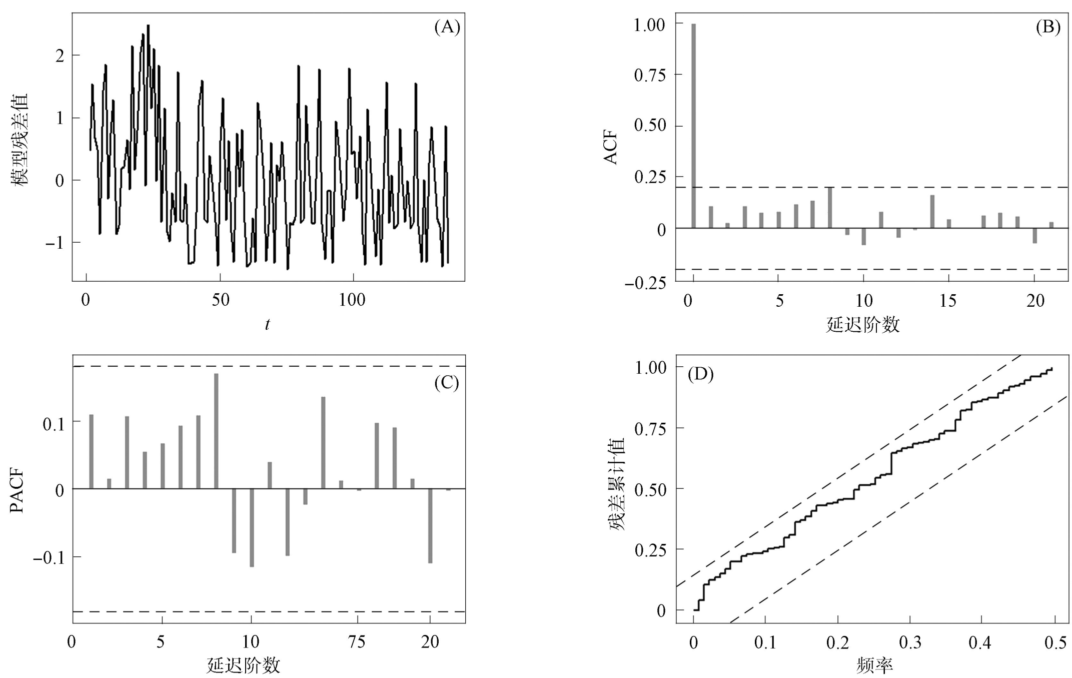

图 4　RCINAR(4)模型拟合残差的路径图(A)，ACF 图(B)，PACF 图(C)和累计周期图(D)

Fig. 4　Path diagram (A), ACF plot (B), PACF plot (C), and cumulative period diagram (D) of fitting residuals of RCINAR(4) model

综上，本文提出了 RCINAR(p)模型，讨论了该模型的统计性质，并通过条件极大似然估计方法估计参数，证明了估计量的渐近正态性，为统计推断提供了可靠依据．为验证模型的有效性和实用性，进行了数值模拟和实例分析．数值模拟结果表明，RCINAR(p)模型能准确捕捉数据特征．在实例分析

中，该模型成功应用于实际案例，进一步证实了其实际应用价值.

参 考 文 献

[1] STEUTEL F W, VAN HARN K. Discrete Analogues of Self-decomposability and Stability [J]. The Annals of Probability, 1979, 7(5): 893-899.

[2] AL-OSH M A, ALZAID A A. First-Order Integer-Valued Autoregressive (INAR(1)) Process [J]. Journal of Time Series Analysis, 1987, 8(3): 261-275.

[3] RISTIĆ M M, BAKOUCH H S, NASTIĆA S. A New Geometric First-Order Integer-Valued Autoregressive (NGINAR(1)) Process [J]. Journal of Statistical Planning and Inference, 2009, 139(7): 2218-2226.

[4] NASTIĆ A S, RISTIĆ M M, BAKOUCH H S. A Combined Geometric INAR(p) Model Based on Negative Binomial Thinning [J]. Mathematical and Computer Modelling, 2012, 55(5/6): 1665-1672.

[5] ZHENG H T, BASAWA I V, DATTA S. First-Order Random Coefficient Integer-Valued Autoregressive Processes [J]. Journal of Statistical Planning and Inference, 2007, 137(1): 212-229.

[6] LIU X F, WANG D H. Estimation of Parameters in the DDRCINAR(p) Model [J]. Brazilian Journal of Probability and Statistics, 2019, 33(3): 638-673.

[7] LIU X F, WANG D H, DENG D L, et al. Maximum Likelihood Estimation of the DDRCINAR(p) Model [J]. Communications in Statistics: Theory and Methods, 2021, 50(24): 6231-6255.

[8] LIU X F, ZHANG W K, YIN W Z, et al. Maximum Likelihood Estimation of the SDMINAR(p) Model to Analyze Some COVID-19 Data [J]. Communications in Statistics: Simulation and Computation, 2024, 53(9): 4307-4330.

[9] LIU X F, YIN W Z, ZHANG W K, et al. A Mixed INAR(p) Model with Serially Dependent Innovation with Application to Some COVID-19 Data [J]. Communications in Statistics: Theory and Methods, 2024, 53(24): 8819-8847.

[10] CHANG L Y, LIU X F, WANG D H, et al. First-Order Random Coefficient Mixed-Thinning Integer-Valued Autoregressive Model [J]. Journal of Computational and Applied Mathematics, 2022, 410: 114222-1-114222-24.

[11] LIU X F, JIANG H, WANG D H. Estimation of Parameters in the MDDRCINAR(p) Model [J]. Journal of Statistical Computation and Simulation, 2023, 93(6): 983-1010.

[12] RISTIĆM M, NASTIĆ A S, MILETIĆILIĆ A V. A Geometric Time Series Model with Dependent Bernoulli Counting Series [J]. Journal of Time Series Analysis, 2013, 34(4): 466-476.

[13] DOUKHAN P, LATOUR A, ORAICHI D. A Simple Integer-Valued Bilinear Time Series Model [J]. Advances in Applied Probability, 2006, 38(2): 559-578.

[14] ZHENG H T, BASAWA I V, DATTA S. Inference for pth-Order Random Coefficient Integer-Valued Autoregressive Processes [J]. Journal of Time Series Analysis, 2006, 27(3): 411-440.

[15] ANDERSON T W, GOODMAN L A. Statistical Inference about Markov Chains [J]. Annals of Mathematical Statistics, 1957, 28(1): 89-110.

[16] FRANKE J, SELIGMANN T H. Conditional Maximum Likelihood Estimates for INAR(1) Processes and Their Application to Modelling Epileptic Seizure Counts [C]//Developments in Time Series Analysis. London: Chapman & Hall, 1993: 310-330.

[17] YANG K, XU N, LI H, et al. Multivariate Threshold Integer-Valued Autoregressive Processes with Explanatory Variables [J]. Applied Mathematical Modelling, 2023, 124: 142-166.

（责任编辑：赵立芹）