

一阶广义零一堆积 Poisson-Lindley 整数值自回归模型的统计推断

张洁, 杨志鹏, 董小刚

(长春工业大学 数学与统计学院, 长春 130012)

摘要: 针对过分散、零一堆积且个体之间具有相依结构的整数值时间序列数据的建模问题, 提出一个具有零一堆积 Poisson-Lindley 新息的一阶广义整数值自回归模型. 首先, 给出模型的一些统计性质: 期望、方差、自协方差和转移概率; 其次, 利用条件极大似然估计方法对模型的未知参数进行估计; 最后, 将该模型应用到一组实际数据中进行拟合, 并用一些评估准则对模型进行验证. 实例分析结果表明, 该模型拟合效果较好.

关键词: 整数值时间序列; 广义二项稀疏算子; 零一堆积 Poisson-Lindley; 条件极大似然估计

中图分类号: O212.1 文献标志码: A 文章编号: 1671-5489(2025)02-0399-12

Statistical Inference for First-Order Generalized Zero-and-One Inflated Poisson-Lindley Integer-Valued Autoregressive Model

ZHANG Jie, YANG Zhipeng, DONG Xiaogang

(School of Mathematics and Statistics, Changchun University of Technology, Changchun 130012, China)

Abstract: Aiming at the modeling problem of overdispersed, zero-and-one inflated integer-valued time series data with interdependent structures between individuals, we proposed a first-order generalized integer-valued autoregressive model with zero-and-one inflated Poisson-Lindley innovation. Firstly, we gave some statistical properties of the model, including expectation, variance, autocovariance, and transition probability. Secondly, the conditional maximum likelihood estimation method was used to estimate the unknown parameters of the model. Finally, the model was applied to a set of real data for fitting, and some evaluation criteria were used to verify the model. The case analysis results show that the model has a good fitting effect.

Keywords: integer-valued time series; generalized binomial thinning operator; zero-and-one inflated Poisson-Lindley; conditional maximum likelihood estimation

时间序列模型是一种用于分析和预测按时间顺序排列的数据或观测值的统计模型. 传统的时间序列模型能很好地处理连续型数据, 但在实际应用中, 经常需要分析一些只能取整数值的数据, 即计数数据. 从随机计数过程中获得的数据称为整数值时间序列, 它在经济学、社会科学、人寿保险等领域

收稿日期: 2024-06-11.

第一作者简介: 张洁(1992—), 女, 汉族, 博士, 讲师, 从事整数值时间序列分析、假设检验和变量选择的研究, E-mail: zhangjie@ccut.edu.cn. **通信作者简介:** 董小刚(1961—), 男, 汉族, 博士, 教授, 从事金融统计分析和生存分析的研究, E-mail: dongxiaogang@ccut.edu.cn.

基金项目: 吉林省自然科学基金自由探索项目(批准号: YDZJ202301ZYTS384).

应用广泛.

Al-Osh 等^[1]在二项稀疏算子^[2]的基础上,建立了一类一阶整数值自回归模型,其中模型以 Poisson 分布为边际,简称为 INAR(1)模型;Alzaid 等^[3]将 INAR(1)模型推广到 p 阶的整数值自回归模型;Al-Osh 等^[4]考虑了 INAR(1)模型中边际分布为几何分布的情形并对其进行了讨论.目前,关于整数值自回归模型及其性质的研究得到广泛关注.但在以 Poisson 分布为边际的整数值自回归模型中,Poisson 分布的均值与方差相等,而实际应用中的计数数据经常表现出过分散等特性,因此应用 Poisson 分布拟合此类数据时会出现一定的局限性.基于此,Sankaran^[5]提出了 Poisson-Lindley (PL) 分布,是一种复合 Poisson 分布,具有单峰性、过分散、无限可分性等性质,在针对具有过分散特征的数据时,相较于一些常规分布具有更好的拟合效果;Mohammadpour 等^[6]提出了具有 Poisson-Lindley 边际的一阶整数值自回归模型,并研究了它的一些统计性质,考虑了几种估计模型未知参数的方法,通过模拟验证比较了这些估计方法的性能.

基于 Poisson-Lindley 分布的整数值自回归模型可以很好地拟合具有过分散类型的数据,但在一些实际问题中,可能还会出现许多零或者一堆积的情况,而基础的模型并不能对其进行有效拟合.针对该类出现堆积情况的数据,Jazi 等^[7]提出了一类带零堆积 Poisson 新息的 INAR(1)模型;Zhu^[8]提出了一类带零堆积 Poisson 和负二项整数值 GARCH 模型;Qi 等^[9]提出了一类具有零和一堆积 Poisson 新息的 INAR(1)模型,即 ZOINAR(1)模型.而考虑到 Poisson-Lindley 分布的优良性能,Xavier 等^[10]提出了零修正 Poisson-Lindley (ZMPL)分布;Sharafi 等^[11]提出了一类具有零修正 Poisson-Lindley 新息的一阶整数值自回归模型,并给出了模型的一些统计性质,采用几种估计方法估计模型未知参数;Mohammadi 等^[12]提出了一类具有零和一堆积 Poisson-Lindley 新息的一阶整数值自回归模型,简称为 ZOIPL-INAR(1)模型. Poisson-Lindley 分布可以有效处理一些实际问题,但当零和一的数量超过 Poisson-Lindley 分布的预期水平时,ZOIPL-INAR(1)模型比 Poisson-Lindley 分布更有效.

二项稀疏算子是基于独立同分布的 Bernoulli 随机变量序列提出的,若此时将模型应用到个体之间具有相依结构的数据中则可能会使该模型的拟合效果不佳.基于此,Ristić等^[13]提出了一个新的算子——广义二项稀疏算子,为具有相依结构的整数值时间序列数据的研究提供了参考,并将广义二项稀疏算子应用于具有几何新息的一阶整数值自回归模型,研究了该模型的一些统计性质,同时给出了参数估计的一些数值模拟结果;Shirozhan 等^[14]基于混合 Pogram 算子和广义二项稀疏算子,提出了一类新的具有几何边际的一阶整数值自回归(MPDBGINAR(1))模型,研究了该模型的统计性质,同时采用基于 EM(expectation maximization)算法的极大似然估计方法估计模型参数,并通过 Monte Carlo 模拟验证了估计值的性能;Kang 等^[15]将广义二项稀疏算子引入到二项自回归模型中,提出了一类一阶广义二项自回归(GBAR(1))模型;在此基础上,Zhang 等^[16-17]分别提出了几类混合 GBAR(1)模型和一类高阶 GBAR 模型.基于上述研究工作,为更有效地刻画个体之间具有相依性且带有零一堆积结构特征的过离散整数值时间序列数据,本文将广义二项稀疏算子引入到 ZOIPL-INAR(1)模型中,建立一类具有零一堆积 Poisson-Lindley 新息的一阶广义整数值自回归模型,并通过参数估计得到有效的拟合结果,为具有相依结构的过分散数据研究提供一定的理论依据.

1 模型理论

1.1 零一堆积 Poisson-Lindley 分布的定义及性质

如果随机变量 Y 的概率质量函数为

$$P(Y=k) = \begin{cases} \varphi_0 + (1 - \varphi_0 - \varphi_1) \frac{\delta^2(\delta+2)}{(1+\delta)^3}, & k=0, \\ \varphi_1 + (1 - \varphi_0 - \varphi_1) \frac{\delta^2(\delta+3)}{(1+\delta)^4}, & k=1, \\ (1 - \varphi_0 - \varphi_1) \frac{\delta^2(\delta+2+k)}{(1+\delta)^{k+3}}, & k=2,3,\dots, \end{cases} \quad (1)$$

则称随机变量 Y 服从零一堆积 Poisson-Lindley 分布^[12], 用 $Y \sim \text{ZOIPL}(\varphi_0, \varphi_1, \delta)$ 表示, 其中 $0 \leq \varphi_0, \varphi_1 < 1, \delta > 0$.

ZOIPL($\varphi_0, \varphi_1, \delta$)分布是零和一堆积分布以及一个 PL(δ)分布的混合, 因此式(1)也可表示为

$$P(Y = k) = \varphi_0 P(\zeta_0 = k) + \varphi_1 P(\zeta_1 = k) + (1 - \varphi_0 - \varphi_1)P(Z = k), \tag{2}$$

其中 $P(\zeta_0 = 0) = P(\zeta_1 = 1) = 1, Z \sim \text{PL}(\delta), P(Z = z) = \frac{\delta^z (z + \delta + 2)}{(\delta + 1)^{z+3}}, z = 0, 1, 2, \dots, \delta > 0$.

特别地, 当 $\varphi_0 = 0$ 时, ZOIPL 分布退化为一个一堆积 Poisson-Lindley (OIPL)分布; 当 $\varphi_1 = 0$ 时, ZOIPL 分布退化为零堆积 Poisson-Lindley (ZIPL)分布; 当 $\varphi_0 = \varphi_1 = 0$ 时, ZOIPL 分布退化为参数为 δ 的标准 Poisson-Lindley 分布. ZOIPL 分布的统计性质如下:

- 1) $E(Y) = \varphi_1 + \varphi_2 \frac{\delta + 2}{\delta(1 + \delta)} = \frac{(1 - \varphi_0)(\delta + 2) + \varphi_1(\delta^2 - 2)}{\delta(1 + \delta)}$;
- 2) $\text{Var}(Y) = \varphi_1(1 - \varphi_1) + \varphi_2 \frac{\delta^3 + 5\delta^2 + 10\delta + 6}{\delta^2(1 + \delta)^2} - \varphi_2^2 \frac{(\delta + 2)^2}{\delta^2(1 + \delta)^2} - 2\varphi_1\varphi_2 \frac{\delta + 2}{\delta(\delta + 1)}$;
- 3) $E(Y^n) = \varphi_1 + \varphi_2 E(Z^n)$.

其中 $\varphi_2 = 1 - \varphi_0 - \varphi_1 \in (0, 1], Z \sim \text{PL}(\delta)$. Mohammadi 等^[12]在注 2.1 中给出了基于 Fisher 离散指数对 ZOIPL 分布离散程度的相关讨论.

1.2 广义二项稀疏算子的定义及性质

广义二项稀疏算子定义^[13]为

$$\alpha \circ_{\theta} X = \sum_{i=1}^X U_i, \tag{3}$$

其中: X 是一个非负的整数值随机变量; $U_i = (1 - V_i)W_i + V_iZ, \{W_i\}$ 是独立同分布的以 α 为参数的 Bernoulli 随机变量序列, $\{V_i\}$ 是独立同分布的以 θ 为参数的 Bernoulli 随机变量序列, Z 为服从参数为 α 的 Bernoulli 分布随机变量; $\alpha \in (0, 1); \theta \in (0, 1)$. 由广义二项稀疏算子的定义可知, U_i 之间的相关性可利用 Z 刻画, 广义二项稀疏算子也可表示为

$$U_1 + U_2 + \dots + U_n \stackrel{d}{=} \begin{cases} \text{Bin}(n, \alpha(1 - \theta)), & \text{w. p. } 1 - \alpha, \\ \text{Bin}(n, \alpha + \theta - \alpha\theta), & \text{w. p. } \alpha, \end{cases} \tag{4}$$

其中 $\text{Bin}(n, \alpha)$ 表示二项分布, w. p. 表示以概率. 广义二项稀疏运算的部分统计性质如下:

- 1) $E(\alpha \circ_{\theta} X | X) = \alpha X$;
- 2) $\text{Var}(\alpha \circ_{\theta} X | X) = \alpha(1 - \alpha)[\theta^2 X^2 + (1 - \theta^2)X]$;
- 3) $\text{Cov}(\alpha \circ_{\theta} X, Y) = \alpha \text{Cov}(X, Y)$.

1.3 模型的定义及性质

下面给出一类具有零一堆积 Poisson-Lindley 新息的一阶广义整数值自回归模型, 简称为 ZOIPL-GINAR(1)模型, 并讨论其基本性质.

定义 1 假设 $\{X_t\}_{t \in \mathbb{N}}$ 是一个 ZOIPL-GINAR(1)过程, 则 $\{X_t\}_{t \in \mathbb{N}}$ 满足

$$X_t = \alpha \circ_{\theta} X_{t-1} + \varepsilon_t, \quad t = 1, 2, \dots, \tag{5}$$

其中: $\alpha \in (0, 1); \circ_{\theta}$ 为式(3)提出的广义二项稀疏算子, $\theta \in (0, 1)$; 序列 $\{\varepsilon_t\}$ 是一个独立同分布的服从 ZOIPL($\varphi_0, \varphi_1, \delta$)分布的随机变量序列, 对所有的 $k \geq 1, \varepsilon_t$ 与 X_{t-k} 均无关, 新息项 ε_t 也与广义二项稀疏算子中包含的计数级数无关.

由于过程 $\{X_t\}$ 是一个 Markov 链, 因此其转移概率有以下形式:

$$\begin{aligned} P_{ij} &= P(X_t = j | X_{t-1} = i) = P(\alpha \circ_{\theta} X_{t-1} + \varepsilon_t = j | X_{t-1} = i) = \\ &= \sum_{k=0}^{\min\{i, j\}} P(\alpha \circ_{\theta} X_{t-1} = k, \varepsilon_t = j - k | X_{t-1} = i) = \\ &= \sum_{k=0}^{\min\{i, j\}} P(\alpha \circ_{\theta} X_{t-1} = k | X_{t-1} = i) \times P(\varepsilon_t = j - k | X_{t-1} = i) = \end{aligned}$$

$$\sum_{k=0}^{\min(i,j)} \left[(1-\alpha) \binom{i}{k} (\alpha(1-\theta))^k (1-\alpha(1-\theta))^{i-k} + \alpha \binom{i}{k} (\theta + \alpha(1-\theta))^k (1-\theta - \alpha(1-\theta))^{i-k} \right] \times P(\epsilon_t = j - k). \quad (6)$$

下面给出 ZOIPL-GINAR(1)过程的严平稳性及遍历性.

命题 1 设序列 $\{X_t\}_{t \in \mathbb{Z}}$ 满足式(5), 则:

- 1) $\{X_t\}_{t \in \mathbb{Z}}$ 是一个不可约、非周期、正常返的 Markov 链;
- 2) $\{X_t\}_{t \in \mathbb{Z}}$ 是遍历的, 并存在一个严平稳过程满足式(5).

证明: 1) $\{X_t\}_{t \in \mathbb{Z}}$ 是 \mathbb{N} 上的 Markov 链, 其转移概率如式(6)所示, 有 $P(X_t = j | X_{t-1} = i) > 0$ 恒成立, 故 $\{X_t\}_{t \in \mathbb{Z}}$ 是不可约、非周期的 Markov 链. 下面证明 $\{X_t\}_{t \in \mathbb{Z}}$ 是正常返的, 只需证明 $\sum_{i=1}^{\infty} P^i(0, 0) = +\infty$, 其中 $P^i(x, y) = P(X_t = y | X_0 = x)$. 已知 $X_0 = 0$, 对式(5)迭代 $(t-1)$ 次可得

$$X_t = \alpha^t \circ_{\theta} X_0 + \sum_{i=0}^{t-1} \alpha^i \circ_{\theta} \epsilon_{t-i}.$$

于是

$$\begin{aligned} P^t(0, 0) &= P(X_t = 0 | X_0 = 0) = P(\alpha^t \circ_{\theta} X_0 = 0 | X_0 = 0) \times P\left(\sum_{i=0}^{t-1} \alpha^i \circ_{\theta} \epsilon_{t-i} = 0\right) = \\ &= P(\epsilon_t = 0) \times \prod_{i=1}^{t-1} P(\alpha^i \circ_{\theta} \epsilon_{t-i} = 0) = \\ &= P(\epsilon_t = 0) \times \prod_{i=1}^{t-1} \sum_{k=0}^{\infty} [P(\alpha^i \circ_{\theta} \epsilon_{t-i} = 0 | \epsilon_{t-i} = k) \times P(\epsilon_{t-i} = k)] = \\ &= P(\epsilon_t = 0) \times \prod_{i=1}^{t-1} \sum_{k=1}^{\infty} [P(\alpha^i \circ_{\theta} \epsilon_{t-i} = 0 | \epsilon_{t-i} = k) \times P(\epsilon_{t-i} = k) + P(\epsilon_{t-i} = 0)] \geq \\ &= P(\epsilon_t = 0) \times \prod_{i=1}^{t-1} P(\epsilon_{t-i} = 0) = \\ &= P(\epsilon_t = 0) P(\epsilon_{t-1} = 0) \cdots P(\epsilon_1 = 0) > 0. \end{aligned}$$

因此 $\sum_{i=1}^{+\infty} P(X_i = 0 | X_0 = 0) = +\infty$, 即证明了 $\{X_t\}_{t \in \mathbb{Z}}$ 是一个正常返的 Markov 链, 因此它是遍历的.

2) 是 1) 的直接结论, 根据文献[18]中定理 1.3 即可证明.

命题 2 假设 $\{X_t\}_{t \in \mathbb{Z}}$ 满足 ZOIPL-GINAR(1)过程, 且 $\mu_{\epsilon} = E(\epsilon_t)$, $\sigma_{\epsilon}^2 = \text{Var}(\epsilon_t)$, 则对于 $t \geq 1$, 有以下结论:

- 1) $E(X_t | X_{t-1} = x) = \alpha x + \mu_{\epsilon}$;
- 2) $\text{Var}(X_t | X_{t-1} = x) = \alpha(1-\alpha)(\theta^2 x^2 + (1-\theta^2)x) + \sigma_{\epsilon}^2$;
- 3) $E(X_t) = \frac{\mu_{\epsilon}}{1-\alpha}$;
- 4) $\text{Var}(X_t) = \frac{1}{1-\alpha(\alpha+\theta^2-\alpha\theta^2)} \times \left[\frac{\alpha\theta^2}{1-\alpha} \mu_{\epsilon}^2 + \alpha(1-\theta^2)\mu_{\epsilon} + \sigma_{\epsilon}^2 \right]$;
- 5) $\gamma(k) = \alpha^k \gamma(0)$.

证明: 结论 1)~3) 和结论 5) 显然成立, 因此只需给出结论 4) 的证明. 由方差公式可知,

$$\begin{aligned} \text{Var}(X_t) &= \text{Var}[E(X_t | X_{t-1})] + E[\text{Var}(X_t | X_{t-1})] = \\ &= \alpha^2 \text{Var}(X_{t-1}) + \alpha(1-\alpha)\theta^2 \text{Var}(X_{t-1}) + \frac{\alpha\theta^2}{1-\alpha} \mu_{\epsilon}^2 + \alpha(1-\theta^2)\mu_{\epsilon} + \sigma_{\epsilon}^2, \end{aligned}$$

进行求和并移项可得

$$[1 - \alpha^2 - \alpha(1-\alpha)\theta^2] \text{Var}(X_t) = \frac{\alpha\theta^2}{1-\alpha} \mu_{\epsilon}^2 + \alpha(1-\theta^2)\mu_{\epsilon} + \sigma_{\epsilon}^2,$$

整理后可得模型的方差.

基于文献[12]中注 2.1 可得以下推论.

推论 1 当新息序列 $\epsilon_t \sim \text{ZOIPL}(\varphi_0, \varphi_1, \delta)$ 是过度离散, 即 $\sigma_\epsilon^2/\mu_\epsilon > 1$ 时, 由式(5)定义的 X_t 的方差-均值比为

$$I_d = \frac{\sigma_X^2}{\mu_X} = \frac{1 - \alpha}{[1 - \alpha(\alpha + \theta^2 - \alpha\theta^2)] \times \mu_\epsilon} \times \left[\frac{\alpha\theta^2}{1 - \alpha} \mu_\epsilon^2 + \alpha(1 - \theta^2)\mu_\epsilon + \sigma_\epsilon^2 \right] = \frac{\alpha\theta^2 \mu_\epsilon + \alpha(1 - \theta^2) + \frac{\sigma_\epsilon^2}{\mu_\epsilon}}{1 + \alpha(1 - \theta^2)} > 1,$$

此时, ZOIPL-GINAR(1)过程也是过度离散的.

2 参数估计

下面讨论 ZOIPL-GINAR(1)模型的参数估计问题. 假设 $\{X_t\}_{t \in \mathbb{N}}$ 是 ZOIPL-GINAR(1)模型的一个严平稳遍历解, 令 $\boldsymbol{\eta} = (\alpha, \varphi_0, \varphi_1, \delta, \theta)^\top$ 为待估参数向量, 考虑通过条件极大似然(CML)方法估计参数 $\boldsymbol{\eta}$.

2.1 条件极大似然估计

对一个固定的初值 x_0 , 模型(5)的条件似然函数为

$$L(\boldsymbol{\eta}) = P(X_1 = x_1, \dots, X_N = x_N | X_0 = x_0) = \prod_{t=1}^N P(X_t = x_t = j | X_{t-1} = x_{t-1} = i), \quad (7)$$

其中转移概率 $P(X_t = j | X_{t-1} = i)$ 由式(6)给出. 则条件极大似然估计量 $\hat{\boldsymbol{\eta}}_{\text{CML}} = (\hat{\alpha}_{\text{CML}}, \hat{\varphi}_0^{\text{CML}}, \hat{\varphi}_1^{\text{CML}}, \hat{\delta}_{\text{CML}}, \hat{\theta}_{\text{CML}})^\top$ 可通过最大化条件对数似然函数得到:

$$l(\boldsymbol{\eta}) = \log(L(\boldsymbol{\eta})) = \sum_{t=1}^N \log(P(X_t | X_{t-1})) = \sum_{t=1}^N \log\left(\sum_{k=0}^{\min\{t,j\}} P(\alpha \circ_\theta X_{t-1} = k | X_{t-1} = i) \times P(\epsilon_t = j - k)\right). \quad (8)$$

下面给出 CML 估计量 $\hat{\boldsymbol{\eta}}_{\text{CML}} = (\hat{\alpha}_{\text{CML}}, \hat{\varphi}_0^{\text{CML}}, \hat{\varphi}_1^{\text{CML}}, \hat{\delta}_{\text{CML}}, \hat{\theta}_{\text{CML}})^\top$ 渐近正态性的相关结果. 假设:

(H₁) 集合 $\{k: P(\epsilon_t = k) = f(k, \boldsymbol{\beta}) > 0\}$ 不依赖于 $\boldsymbol{\beta} = (\varphi_0, \varphi_1, \delta)^\top \triangleq (\beta_1, \beta_2, \beta_3)^\top$;

(H₂) $E(\epsilon_t^3) = \varphi_1 + \varphi_2 \left(\frac{\delta + 2}{\delta(1 + \delta)}\right)^3 < \infty$;

(H₃) $P(\epsilon_t = k)$ 关于 $\boldsymbol{\beta}$ 三阶连续可微;

(H₄) 对任意的 $\boldsymbol{\beta}' \in B$, 存在 $\boldsymbol{\beta}'$ 的一个邻域 U , 使得

$$\sum_{k=0}^{\infty} \sup_{\boldsymbol{\beta} \in U} f(k, \boldsymbol{\beta}) < \infty, \quad \sum_{k=0}^{\infty} \sup_{\boldsymbol{\beta} \in U} \left| \frac{\partial f(k, \boldsymbol{\beta})}{\partial \beta_u} \right| < \infty, \quad u = 1, 2, 3,$$

$$\sum_{k=0}^{\infty} \sup_{\boldsymbol{\beta} \in U} \left| \frac{\partial^2 f(k, \boldsymbol{\beta})}{\partial \beta_u \partial \beta_v} \right| < \infty, \quad u, v = 1, 2, 3,$$

其中 B 是 $[0, 1) \times [0, 1) \times (0, +\infty)$ 上的一个开子集;

(H₅) 对任意的 $\boldsymbol{\beta}' \in B$, 存在一个 $\boldsymbol{\beta}'$ 的邻域 U 和递增序列 $\psi_u(j), \psi_{uv}(j), \psi_{uvw}(j) (u, v, w = 1, 2, 3, j \geq 0)$, 使得对所有的 $\boldsymbol{\beta} \in U, k \leq j$ 及 $f(k, \boldsymbol{\beta})$, 均有

$$\left| \frac{\partial f(k, \boldsymbol{\beta})}{\partial \beta_u} \right| \leq \psi_u(j) f(k, \boldsymbol{\beta}), \quad \left| \frac{\partial f(k, \boldsymbol{\beta})}{\partial \beta_u \partial \beta_v} \right| \leq \psi_{uv}(j) f(k, \boldsymbol{\beta}), \quad \left| \frac{\partial f(k, \boldsymbol{\beta})}{\partial \beta_u \partial \beta_v \partial \beta_w} \right| \leq \psi_{uvw}(j) f(k, \boldsymbol{\beta}),$$

并且 $\{X_t\}$ 是平稳分布, 故 $E\psi_u(X_1) < \infty, EX_1\psi_{uv}(X_2) < \infty, E\psi_u(X_1)\psi_{uv}(X_2) < \infty, E\psi_{uvw}(X_2) < \infty$;

(H₆) 令 $\mathbf{I}(\boldsymbol{\eta}) = (\sigma_{pq}(\boldsymbol{\eta}))_{5 \times 5}$ 表示 Fisher 信息矩阵, 即

$$\sigma_{11}(\boldsymbol{\eta}) = E\left(\frac{\partial}{\partial \alpha} \log P(X_1, X_2)\right)^2,$$

$$\sigma_{1p}(\boldsymbol{\eta}) = E\left(\frac{\partial}{\partial \alpha} \log P(X_1, X_2) \frac{\partial}{\partial \beta_{(p-1)}} \log P(X_1, X_2)\right) = \sigma_{p1}(\boldsymbol{\eta}), \quad p = 2, 3, 4,$$

$$\sigma_{pq}(\boldsymbol{\eta}) = E\left(\frac{\partial}{\partial \beta_{(p-1)}} \log P(X_1, X_2) \frac{\partial}{\partial \beta_{(q-1)}} \log P(X_1, X_2)\right), \quad p, q = 2, 3, 4,$$

$$\sigma_{\rho p}(\boldsymbol{\eta}) = E\left(\frac{\partial}{\partial \theta} \log P(X_1, X_2) \frac{\partial}{\partial \beta_{(p-1)}} \log P(X_1, X_2)\right) = \sigma_{\rho 5}(\boldsymbol{\eta}), \quad p = 2, 3, 4,$$

$$\sigma_{55}(\boldsymbol{\eta}) = E\left(\frac{\partial}{\partial \theta} \log P(X_1, X_2)\right)^2,$$

其中: $\mathbf{I}(\boldsymbol{\eta})$ 是非奇异矩阵; $P(X_1, X_2)$ 表示转移概率, 由式(6)给出.

定理 1 在假设条件 $(H_1) \sim (H_6)$ 成立的情况下, 条件极大似然估计 $\hat{\boldsymbol{\eta}}_{\text{CML}}$ 是相合的, 并有如下渐近分布:

$$\sqrt{N}(\hat{\boldsymbol{\eta}}_{\text{CML}} - \boldsymbol{\eta}) \xrightarrow{d} N(0, \mathbf{I}^{-1}(\boldsymbol{\eta})), \quad (9)$$

其中 $\mathbf{I}(\boldsymbol{\eta})$ 是 Fisher 信息矩阵.

证明: 与文献[19]中定理 3.3.1 和定理 3.3.2 及文献[20]中定理 3.2 的讨论类似, 由于定理 1 是文献[21]中定理 2.1 和定理 2.2 的特例, 因此只需证明假设条件 $(H_1) \sim (H_6)$ 成立, 即可得文献[21]中定理 2.1 和定理 2.2 成立.

条件 $(H_1) \sim (H_4)$ 显然满足, 此外, 由零一堆积 Poisson-Lindley 分布的性质及 ZOIPL-GINAR(1) 过程的期望和方差可知, 条件 (H_5) 也成立:

$$\begin{aligned} \psi_1(j) &= c, \quad \psi_2(j) = c, \quad \psi_3(j) = c_1(c_2 + j), \\ \psi_{13}(j) &= cj, \quad \psi_{23}(j) = cj, \quad \psi_{31}(j) = c_1(c_2 + j), \quad \psi_{32}(j) = c_1(c_2 + j), \\ \psi_{33}(j) &= c_1(c_2 + j^2), \quad \psi_{313}(j) = cj^2, \quad \psi_{323}(j) = cj^2, \quad \psi_{333}(j) = cj^3, \end{aligned}$$

其他值都为零, c, c_1, c_2 是一些合适的常数. 对于条件 (H_6) , 只需证明以下条件成立:

$$\begin{aligned} E\left|\frac{\partial}{\partial \alpha} \log P(X_1, X_2)\right|^2 &< \infty, \quad E\left|\frac{\partial}{\partial \beta_u} \log P(X_1, X_2)\right|^2 < \infty, \quad E\left|\frac{\partial}{\partial \theta} \log P(X_1, X_2)\right|^2 < \infty, \\ E\left|\frac{\partial}{\partial \alpha} \log P(X_1, X_2) \cdot \frac{\partial}{\partial \beta_u} \log P(X_1, X_2)\right| &< \infty, \quad E\left|\frac{\partial}{\partial \beta_u} \log P(X_1, X_2) \cdot \frac{\partial}{\partial \theta} \log P(X_1, X_2)\right| < \infty, \\ E\left|\frac{\partial}{\partial \alpha} \log P(X_1, X_2) \cdot \frac{\partial}{\partial \theta} \log P(X_1, X_2)\right| &< \infty, \\ E\left|\frac{\partial}{\partial \alpha} \log P(X_1, X_2) \cdot \frac{\partial}{\partial \beta_u} \log P(X_1, X_2) \cdot \frac{\partial}{\partial \theta} \log P(X_1, X_2)\right| &< \infty. \end{aligned}$$

下面证明转移概率 $P(X_1, X_2)$ 取对数再对参数 α 求偏导, 即 $\frac{\partial}{\partial \alpha} \log P(X_1, X_2)$ 的二阶矩有界. 根据转移概率可得

$$\begin{aligned} P(X_1, X_2) &= \sum_{k=0}^{\min(X_1, X_2)} \left[(1-\alpha) \binom{X_1}{k} (\alpha(1-\theta))^k (1-\alpha(1-\theta))^{X_1-k} + \right. \\ &\quad \left. \alpha \binom{X_1}{k} (\theta + \alpha(1-\theta))^k (1-\theta - \alpha(1-\theta))^{X_1-k} \right] P(\epsilon_t = X_2 - k). \end{aligned}$$

令 $C_1 = (1-\alpha)(\alpha(1-\theta))^k (1-\alpha(1-\theta))^{X_1-k}$, $C_2 = \alpha(\theta + \alpha(1-\theta))^k (1-\theta - \alpha(1-\theta))^{X_1-k}$, 则

$$P(X_1, X_2) = \sum_{k=0}^{\min(X_1, X_2)} \binom{X_1}{k} (C_1 + C_2) P(\epsilon_t = X_2 - k).$$

故 $\frac{\partial}{\partial \alpha} \log P(X_1, X_2) = \frac{1}{P(X_1, X_2)} \frac{\partial P(X_1, X_2)}{\partial \alpha}$, 其中

$$\begin{aligned} \frac{\partial P(X_1, X_2)}{\partial \alpha} &= \sum_{k=0}^{\min(X_1, X_2)} \left[\binom{X_1}{k} C_1 \left[-\frac{1}{1-\alpha} + \frac{k}{\alpha} - \frac{(1-\theta)(X_1-k)}{(1-\alpha(1-\theta))} \right] + \right. \\ &\quad \left. \binom{X_1}{k} C_2 \left[\frac{1}{\alpha} + \frac{k(1-\theta)}{\theta + \alpha(1-\theta)} - \frac{(X_1-k)}{(1-\alpha)} \right] \right] P(\epsilon_t = X_2 - k) = \\ &= \sum_{k=0}^{\min(X_1, X_2)} \left[\binom{X_1}{k} \left[C_1 \left[-\frac{1}{1-\alpha} + \frac{k}{\alpha} - \frac{(1-\theta)(X_1-k)}{(1-\alpha(1-\theta))} \right] + \right. \right. \\ &\quad \left. \left. C_2 \left[\frac{1}{\alpha} + \frac{k(1-\theta)}{\theta + \alpha(1-\theta)} - \frac{(X_1-k)}{(1-\alpha)} \right] \right] \right] P(\epsilon_t = X_2 - k) \leq \end{aligned}$$

$$\begin{aligned} & \sum_{k=0}^{\min\{X_1, X_2\}} \binom{X_1}{k} \left(\frac{k}{\alpha} C_1 + \left(\frac{(k+1)(1-\theta) + \theta/\alpha}{\theta + \alpha(1-\theta)} \right) C_2 \right) P(\epsilon_t = X_2 - k) \leq \\ & \left(\frac{m}{\alpha} + \frac{(m+1)(1-\theta) + \theta/\alpha}{\theta + \alpha(1-\theta)} \right) \sum_{k=0}^{\min\{X_1, X_2\}} \binom{X_1}{k} (C_1 + C_2) P(\epsilon_t = X_2 - k) = \\ & \left(\frac{m}{\alpha} + \frac{(m+1)(1-\theta) + \theta/\alpha}{\theta + \alpha(1-\theta)} \right) P(X_1, X_2) \leq \frac{2(m+1)}{\alpha} P(X_1, X_2), \end{aligned}$$

同理, 可得 $\frac{\partial P(X_1, X_2)}{\partial \alpha} \geq -\frac{X_1+1}{1-\alpha}$, 其中 $m = \min\{X_1, X_2\}$, 故可得

$$-\frac{X_1+1}{1-\alpha} \leq \frac{\partial}{\partial \alpha} \log P(X_1, X_2) \leq \frac{2(X_1+1)}{\alpha} \tag{10}$$

对所有满足 $P(X_1, X_2) > 0$ 和 $\alpha > 0$ 的 X_1, X_2 均成立, 表明在平稳状态下,

$$E \left| \frac{\partial}{\partial \alpha} \log P(X_1, X_2) \right|^2 \leq c_1 \cdot E(X_1 + c_2)^2 < \infty \tag{11}$$

成立, 其中 c_1, c_2 是一个合适的常数. 类似地, 由条件(H₅)可得

$$\left| \frac{\partial}{\partial \beta_u} \log P(X_1, X_2) \right| \leq \frac{1}{P(X_1, X_2)} \sum_{k=0}^{\min\{X_1, X_2\}} \binom{X_1}{k} (C_1 + C_2) \left| \frac{\partial P(X_1, X_2)}{\partial \beta_u} \right| \leq \psi_u(X_2),$$

因此

$$E \left| \frac{\partial}{\partial \beta_u} \log P(X_1, X_2) \right|^2 \leq E(\psi_u(X_2))^2 < \infty. \tag{12}$$

同理可得

$$-\frac{X_1}{\alpha(1-\theta)} \leq \frac{\partial}{\partial \theta} \log P(X_1, X_2) \leq \frac{X_1}{\theta + \alpha},$$

故得到对所有满足 $P(X_1, X_2) > 0$ 和 $\theta > 0$ 的 X_1, X_2 也均成立, 表明在平稳状态下,

$$E \left| \frac{\partial}{\partial \theta} \log P(X_1, X_2) \right|^2 \leq C \cdot E(X_1)^2 < \infty \tag{13}$$

同样成立, 其中 C 是一个合适的常数. 由式(11)~(13)和条件(H₅), 可以证明

$$E \left| \frac{\partial}{\partial \alpha} \log P(X_1, X_2) \cdot \frac{\partial}{\partial \beta_u} \log P(X_1, X_2) \right| < \infty,$$

$$E \left| \frac{\partial}{\partial \beta_u} \log P(X_1, X_2) \cdot \frac{\partial}{\partial \theta} \log P(X_1, X_2) \right| < \infty,$$

$$E \left| \frac{\partial}{\partial \alpha} \log P(X_1, X_2) \cdot \frac{\partial}{\partial \theta} \log P(X_1, X_2) \right| < \infty,$$

$$E \left| \frac{\partial}{\partial \alpha} \log P(X_1, X_2) \cdot \frac{\partial}{\partial \beta_u} \log P(X_1, X_2) \cdot \frac{\partial}{\partial \theta} \log P(X_1, X_2) \right| < \infty$$

均成立. 因此 Fisher 信息矩阵 $\mathbf{I}(\boldsymbol{\eta})$ 有意义. 最后通过一些简单且繁琐的计算表明条件(H₆)也成立.

2.2 数值模拟

下面对 ZOIPL-GINAR(1)模型进行数值模拟, 通过参数估计的偏差和均方误差比较条件极大似然估计的估计效果. 根据参数 $\alpha, \varphi_0, \varphi_1, \delta, \theta$ 的取值空间, 选取以下 4 组不同的参数组合进行数值模拟.

第一组参数: $\alpha=0.4, \varphi_0=0.4, \varphi_1=0.4, \delta=0.5, \theta=0.4$;

第二组参数: $\alpha=0.3, \varphi_0=0.5, \varphi_1=0.3, \delta=0.5, \theta=0.4$;

第三组参数: $\alpha=0.5, \varphi_0=0.5, \varphi_1=0.3, \delta=0.7, \theta=0.6$;

第四组参数: $\alpha=0.3, \varphi_0=0.4, \varphi_1=0.4, \delta=0.7, \theta=0.6$.

ZOIPL-GINAR(1)模型在上述 4 组参数下的样本路径和自相关函数(ACF)如图 1 所示, 其中样本量 $N=100$. 由图 1 可见, 4 个序列都是平稳的且表现出一定的零一堆积特征.

对上述 4 组参数组合, 在 R 软件环境下进行 1 000 次重复实验, 分别在样本量 N 为 100, 200, 500,

1 000 的 4 种情形下给出估计结果. 表 1 列出了在不同参数组合下模型的条件极大似然估计结果. 由表 1 可见, 随着样本量的增大, 估计值的偏差(Bias)和均方误差(MSE)均越来越小, 表明模型参数的条件极大似然估计值是相合且有效的, 因此用条件极大似然估计方法可以对模型参数进行可靠估计.

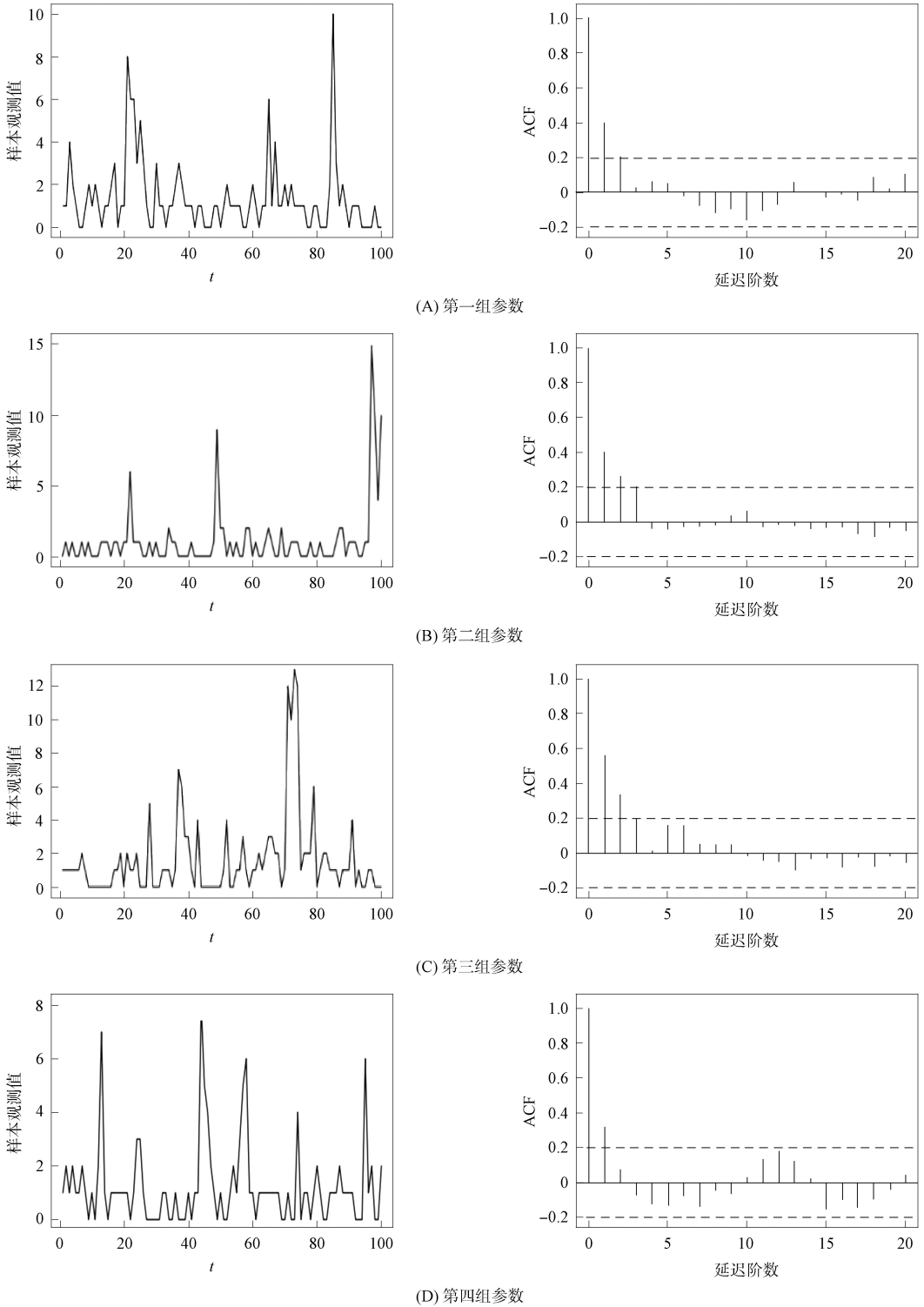


图 1 4 组参数的样本路径图和 ACF 图

Fig. 1 Sample path diagrams and ACF diagrams of four sets of parameters

图 2 为当样本量为 500 时第一组参数下 ZOIPL-GINAR(1)模型的条件极大似然估计值的 Q-Q 图. 由图 2 可见, 其曲线大致呈预期的正态分布. 采用其他参数组合的 ZOIPL-GINAR(1)模型也可得

到类似结果. 图 3 为当样本量 $N=100, 200, 500, 1\ 000$ 时在第一组参数下模型条件极大似然估计值的箱线图. 由图 3 可见, 随着样本量的增大, 估计值的 Bias 和 MSE 均减小.

表 1 不同样本量下 4 组参数的条件极大似然估计结果

Table 1 Conditional maximum likelihood estimation results of four sets of parameters under different sample sizes

组别	参数	N=100		N=200		N=500		N=1 000	
		Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
第一组	$\alpha=0.4$	-0.013 3	0.008 5	-0.006 3	0.005 2	-0.003 6	0.003 2	-0.001 2	0.002 6
	$\varphi_0=0.4$	-0.019 2	0.012 7	-0.011 5	0.006 9	-0.006 3	0.003 9	-0.001 7	0.003 0
	$\varphi_1=0.4$	-0.005 4	0.009 6	-0.001 3	0.005 3	0.001 1	0.003 3	0.000 2	0.002 5
	$\delta=0.5$	0.064 9	0.078 8	0.025 6	0.025 7	0.010 1	0.014 5	0.004 7	0.011 1
	$\theta=0.4$	-0.070 0	0.051 9	-0.032 8	0.024 9	-0.009 7	0.008 9	-0.005 4	0.004 5
第二组	$\alpha=0.3$	-0.011 3	0.026 6	-0.006 6	0.022 1	-0.001 7	0.019 5	-0.000 5	0.018 8
	$\varphi_0=0.5$	-0.016 7	0.022 1	-0.005 8	0.020 4	-0.003 6	0.018 4	-0.002 4	0.018 1
	$\varphi_1=0.3$	-0.003 0	0.023 7	-0.003 9	0.021 2	-0.000 7	0.019 1	0.000 4	0.018 4
	$\delta=0.5$	0.062 1	0.082 6	0.023 0	0.035 5	0.010 5	0.024 7	0.005 1	0.020 9
	$\theta=0.4$	-0.075 5	0.068 5	-0.044 5	0.041 7	-0.018 7	0.018 7	-0.004 6	0.011 7
第三组	$\alpha=0.5$	-0.016 7	0.025 6	-0.010 5	0.022 5	-0.002 7	0.019 4	-0.002 6	0.018 7
	$\varphi_0=0.5$	-0.051 8	0.061 7	-0.020 6	0.024 9	-0.007 9	0.019 3	-0.003 4	0.019 0
	$\varphi_1=0.3$	-0.009 0	0.079 0	-0.001 7	0.069 9	-0.001 0	0.067 9	-0.000 5	0.066 7
	$\delta=0.7$	0.139 6	0.366 8	0.057 1	0.130 2	0.021 9	0.076 9	0.008 9	0.060 0
	$\theta=0.6$	-0.033 4	0.054 5	-0.017 0	0.034 6	-0.003 9	0.027 3	-0.000 6	0.025 4
第四组	$\alpha=0.3$	-0.012 8	0.066 0	-0.005 4	0.059 5	-0.002 5	0.056 2	-0.001 7	0.055 2
	$\varphi_0=0.4$	-0.046 1	0.065 6	-0.017 3	0.035 3	-0.005 4	0.030 4	-0.003 6	0.029 3
	$\varphi_1=0.4$	-0.016 8	0.039 1	-0.009 1	0.032 5	-0.002 3	0.029 1	-0.000 2	0.028 7
	$\delta=0.7$	0.146 9	0.351 2	0.063 5	0.156 4	0.021 4	0.094 5	0.011 4	0.082 1
	$\theta=0.6$	-0.072 3	0.102 8	-0.029 5	0.056 6	-0.009 3	0.043 3	-0.004 9	0.037 6

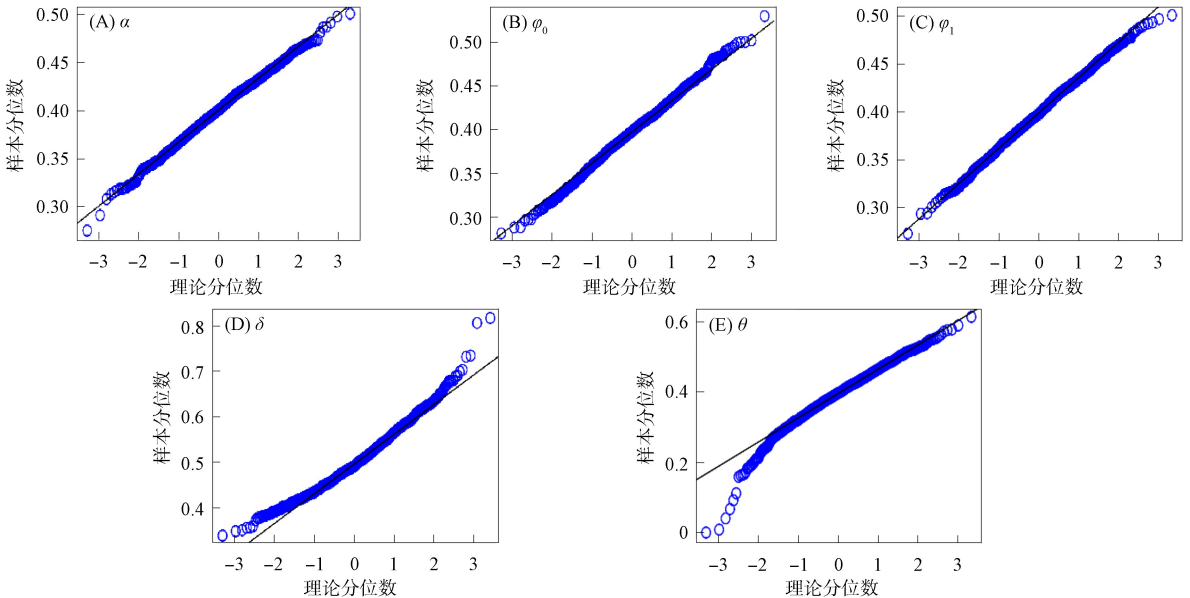


图 2 当样本量 $N=500$ 时在第一组参数下 CML 估计量的 Q-Q 图

Fig. 2 Q-Q plots of CML estimators under the first set of parameters when sample size $N=500$

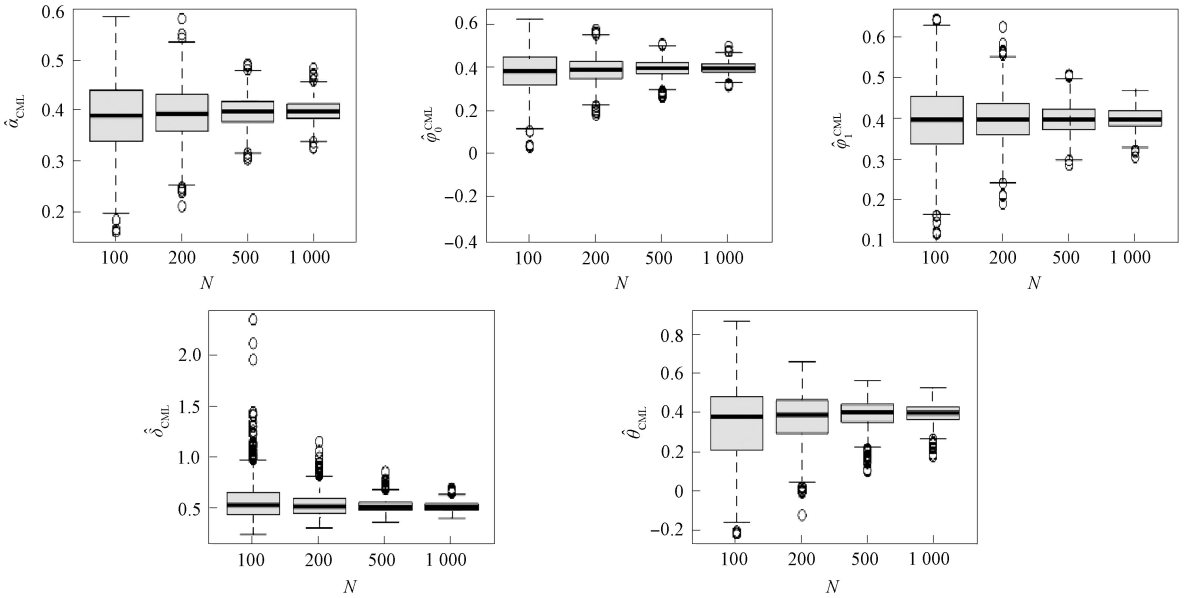


图 3 当样本量 $N=100, 200, 500, 1\ 000$ 时在第一组参数下 CML 估计量的箱线图

Fig. 3 Box plots of CML estimators under the first set of parameters when sample sizes $N=100, 200, 500, 1\ 000$

3 实例分析

实例分析数据采用 Kang 等^[22]给出的在 2006 年 6 月 1 日至 2007 年 2 月 28 日期间, CW β TeXpert 程序的每日下载次数数据. 该组数据由 267 个观测值组成, 其中 74 个(占总量的 27.72%)观测值为 0, 57 个(占总量的 21.35%)观测值为 1, 其余 136 个(占总量的 50.93%)观测值为该程序的每日下载次数 1 次以上的数量. 通过离散指数 $I_d = \text{Var}(Y)/E(Y)$ 测量随机变量 Y 的离散程度, 如果 $I_d > 1 (< 1)$ 或 $I_d = 1$, 则随机变量 Y 分别表现为过离散(欠离散)或等离散. 利用公式计算该数据集的离散指数为 3.138 3, 说明该数据集过离散, 且由推论 1 可知 ZOIPL-GINAR(1) 过程可刻画过离散数据, 因此可考虑使用 ZOIPL-GINAR(1) 模型进行拟合. 原始数据样本路径图及其 ACF 图和 PACF 图如图 4 所示,

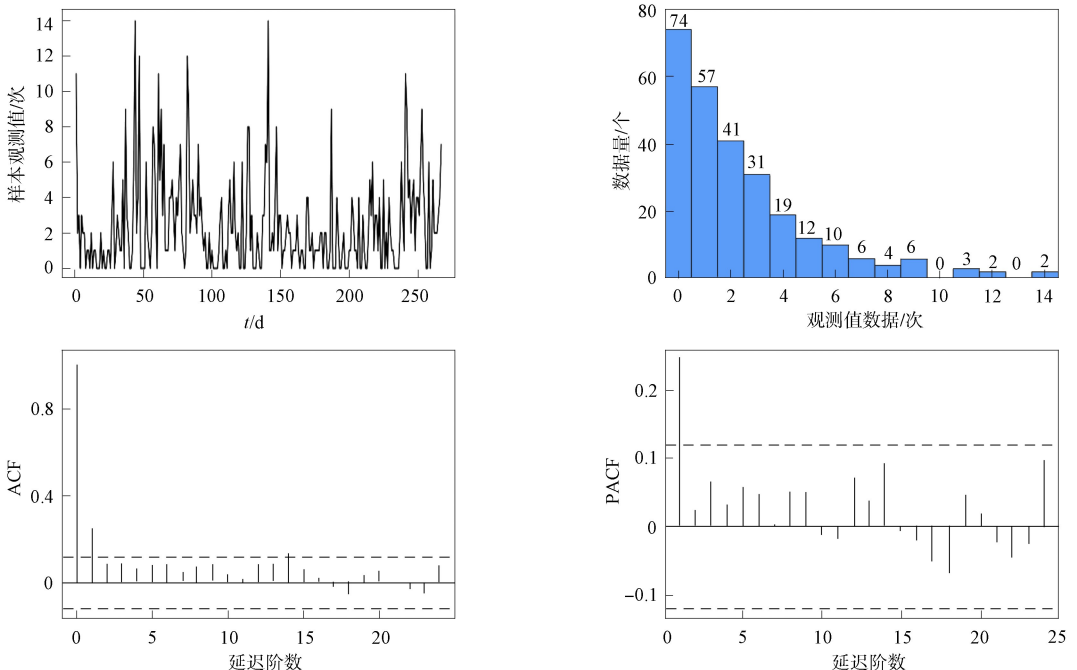


图 4 2006-06-01—2007-02-28 期间, CW β TeXpert 程序的每日下载次数数据集信息

Fig. 4 Daily download count dataset information of CW β TeXpert program during period of 2006-06-01—2007-02-28

该组数据的最小值为 0.000, 最大值为 14.000, 中位数为 2.000, 均值为 2.401, 方差为 7.534, 离散指数为 3.138 3, 0 占比为 27.72%, 1 占比 21.35%。将 ZOIPL-GINAR(1)模型应用于该组数据, 并与 ZOIPL-INAR(1), ZIPL-INAR(1), OIPL-INAR(1), PL-INAR(1), ZINAR(1), NGINAR(1) 和 INAR(1)模型进行比较, 通过比较各模型参数的赤池信息准则(AIC)和离散指数(I_d)说明 ZOIPL-GINAR(1)模型拟合的优良性。实例分析结果列于表 2。由表 2 的 AIC 和 I_d 值可见, ZOIPL-GINAR(1)模型比其他模型拟合效果更好。

表 2 不同模型实例分析结果比较

Table 2 Comparison of case analysis results by different models

模型	$\hat{\alpha}$	$\hat{\varphi}_0$	$\hat{\varphi}_1$	$\hat{\delta}$	$\hat{\lambda}$	$\hat{\mu}$	$\hat{\theta}$	AIC	I_d
ZOIPL-GINAR(1)	0.255 8	0.197 2	0.061 4	0.688 6			0.593 2	1 078.609	3.289 7
ZOIPL-INAR(1)	0.158 5	0.173 6	0.062 0	0.638 3				1 083.571	2.895 3
ZIPL-INAR(1)	0.156 7	0.137 4		0.689 0				1 083.186	2.719 4
OIPL-INAR(1)	0.119 4		0.010 0	0.760 0				1 088.058	2.340 6
PL-INAR(1)	0.117 9			0.755 5				1 086.113	2.530 2
ZINAR(1)	0.242 8	0.429 0			3.134 0			1 143.534	2.082 0
NGINAR(1)	0.151 4					2.330 9		1 082.431	3.330 9
INAR(1)	0.163 0				2.013 8			1 296.098	1.000 0

下面利用

$$e_t = \frac{X_t - E(X_t | X_{t-1})}{\sqrt{\text{Var}(X_t | X_{t-1})}}, \quad t = 2, 3, \dots, N \tag{14}$$

对 ZOIPL-GINAR(1)模型的 Pearson 残差进行分析。根据 ZOIPL-GINAR(1)模型的条件期望和条件方差公式, 计算该模型在条件极大似然估计方法下的残差均值和方差。估计拟合残差的均值和方差分别越接近 0 和 1, 模型的拟合效果越好。在实际数据中, 条件极大似然估计方法下所拟合的 ZOIPL-GINAR(1)模型残差的均值和方差分别为 0.009 9 和 1.139 2。该结果再次说明了用 ZOIPL-GINAR(1)模型拟合这组数据更合适。

综上所述, 将 ZOIPL-GINAR(1)模型应用于 CW β TeXpert 程序的每日下载次数数据进行拟合, 与几类整数值自回归模型进行比较的结果表明, ZOIPL-GINAR(1)模型的拟合效果优于其他模型, 且该模型不仅能拟合存在过分散以及零和一堆积的数据, 还能刻画数据中个体之间存在的相依情况。

参 考 文 献

[1] 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.

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

[3] ALZAID A A, AL-OSH M A. An Integer-Valued p th-Order Autoregressive Structure (INAR(p)) Process [J]. Journal of Applied Probability, 1990, 27(2): 314-324.

[4] AL-OSH M A, ALY E A. First Order Autoregressive Time Series with Negative Binomial and Geometric Marginals [J]. Communication in Statistics: Theory and Methods, 1992, 21(9): 2483-2492.

[5] SANKARAN M. The Discrete Poisson-Lindley Distribution [J]. Biometrics, 1970, 26: 145-149.

[6] MOHAMMADPOUR M, BAKOUCH H S, SHIROZHAN M. Poisson-Lindley INAR(1) Model with Applications [J]. Brazilian Journal of Probability and Statistics, 2018, 32(2): 262-280.

[7] JAZI M A, JONES G, LAI C. First-Order Integer Valued AR Processes with Zero Inflated Poisson Innovations [J]. Journal of Time Series Analysis, 2012, 33(6): 954-963.

[8] ZHU F K. Zero-Inflated Poisson and Negative Binomial Integer-Valued GARCH Models [J]. Journal of Statistical Planning and Inference, 2012, 142(4): 826-839.

- [9] QI X H, LI Q, ZHU F K. Modeling Time Series of Count with Excess Zeros and Ones Based on INAR(1) Model with Zero-and-One Inflated Poisson Innovations [J]. *Journal of Computational and Applied Mathematics*, 2019, 346: 572-590.
- [10] XAVIER D, SANTOS-NETO M, BOURGUIGNON M, et al. Zero-Modified Poisson-Lindley Distribution with Applications in Zero-Inflated and Zero-Deflated Count Data [EB/OL]. (2017-12-12)[2024-03-19]. <https://doi.org/10.48550/arxiv.1712.04088>.
- [11] SHARAFI M, SAJJADNIA Z, ZAMANI A. A First-Order Integer-Valued Autoregressive Process with Zero-Modified Poisson-Lindley Distributed Innovations [J]. *Communications in Statistics: Simulation and Computation*, 2023, 52(3): 685-702.
- [12] MOHAMMADI Z, SAJJADNIA Z, BAKOUCH H S, et al. Zero-and-One Inflated Poisson-Lindley INAR(1) Process for Modelling Count Time Series with Extra Zeros and Ones [J]. *Journal of Statistical Computation and Simulation*, 2022, 92(10): 2018-2040.
- [13] RISTIĆ M M, NASTIĆ A S, ILIĆ A V M. A Geometric Time Series Model with Dependent Bernoulli Counting Series [J]. *Journal of Time Series Analysis*, 2013, 34: 466-476.
- [14] SHIROZHAN M, MOHAMMADPOUR M, BAKOUCH H S. A New Geometric INAR(1) Model with Mixing Pogram and Generalized Binomial Thinning Operators [J]. *Iranian Journal of Science and Technology, Transactions A: Science*, 2019, 43(3): 1011-1020.
- [15] KANG Y, WANG D H, YANG K. Extended Binomial AR(1) Process with Generalized Binomial Thinning Operator [J]. *Communications in Statistics: Theory and Methods*, 2020, 49(14): 3498-3520.
- [16] ZHANG J, SHAO S Y, YANG K, et al. A Statistical Study for Some Classes of First-Order Mixed Generalized Binomial Autoregressive Models [J]. *Communications in Statistics: Theory and Methods*, 2024, 53(14): 5057-5075.
- [17] ZHANG J, SHAO S Y, WANG D H, et al. Statistical Inference of p th-Order Generalized Binomial Autoregressive Model [J]. *Journal of the Korean Statistical Society*, 2024, 53(4): 1003-1026.
- [18] KARLIN S, TAYLOR H E. *A First Course in Stochastic Processes* [M]. 2nd ed. New York: Academic Press, 1975: 1-576.
- [19] 齐霄虹. 基于零一膨胀的泊松 INAR(1)模型 [D]. 长春: 吉林大学, 2018. (QI X H. INAR(1) Model Based on Zero-and-One Inflated Poisson Innovations [D]. Changchun: Jilin University, 2018.)
- [20] YANG K, LI H, WANG D H, et al. Random Coefficients Integer-Valued Threshold Autoregressive Processes Driven by Logistic Regression [J]. *Asta-Advances in Statistical Analysis*, 2021, 105(4): 533-557.
- [21] BILLINGSLEY P. *Statistical Inference for Markov Processes* [M]. Chicago: University of Chicago Press, 1961: 10-14.
- [22] KANG Y, ZHU F K, WANG D H, et al. A Zero-Modified Geometric INAR(1) Model for Analyzing Count Time Series with Multiple Features [J]. *The Canadian Journal of Statistics*, 2023, 52(3): 873-899.

(责任编辑: 李琦)