

Po-RCMTINAR(1)模型的矩估计

陈 瑜, 刘秀芳*, 赵丽华

太原理工大学数学学院, 山西 太原

Email: 1121931963@qq.com, *liuxiufang@tyut.edu.cn, zlh5259@163.com

收稿日期: 2021年5月24日; 录用日期: 2021年6月13日; 发布日期: 2021年6月25日

摘 要

本文提出了一阶泊松随机系数混合算子整数值自回归时间序列模型(Po-RCMTINAR(1))。对该模型的矩和自协方差函数进行了研究, 并使用Yule-Walker方法估计估计模型中的未知参数。最后, 分析了一组实际数据集并与MTINAR(1)模型比较, 根据MSE准则得到一阶泊松随机系数混合算子整数值自回归模型(Po-RCMTINAR(1))更适合分析此数据集。

关键词

Po-RCMTINAR(1)模型, Yule-Walker估计, MSE准则

Moment Estimation of the Po-RCMTINAR(1) Model

Yu Chen, Xiufang Liu*, Lihua Zhao

College of Mathematics, Taiyuan University of Technology, Taiyuan Shanxi

Email: 1121931963@qq.com, *liuxiufang@tyut.edu.cn, zlh5259@163.com

Received: May 24th, 2021; accepted: Jun. 13th, 2021; published: Jun. 25th, 2021

Abstract

This paper proposes a first-order random coefficient mixed thinning integer-valued autoregressive time series model (Po-RCMTINAR(1)) as the innovation sequence has a Poisson distribution. The moment and auto-covariance function of the model are studied, and the unknown parameters are estimated using the Yule-Walker method. Finally, a set of actual data sets are used to compare with the MTINAR(1) model. According to the MSE criterion, the first-order Poisson random coefficient mixed thinning integer-valued autoregressive model (Po-RCMTINAR (1)) is more suitable for analyzing this data set.

*通讯作者。

Keywords

Po-RCMTINAR(1) Model, Yule-Walker Estimation, MSE Criterion

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在现实生活中, 整数值时间序列数据非常常见, 它是由某种现象的某一统计指标在不同时间上的状态所形成的整数计数数据。这种类型的数据广泛应用于通讯保障、医疗卫生、法律、保险精算等众多领域, 例如某地区某医院某种疾病每月住院的病人数, 某地区某月的事故发生数等[1]。因此, 对整值时间序列的研究受到越来越多学者的关注。大部分这一类数据通常显示出短期的相依关系且是取非负整数值的, 使用一般的时间序列模型来拟合这些数据会引起异常预测。所以相较于传统的取连续实值得时间序列数据, 研究整值数据更加困难。

1979年, Steutal 和 Van Harn 提出了稀疏算子[2], 1987年, Al-Osh 和 Alzaid 基于二项稀疏算子提出了一阶整数值自回归模型[3], 该模型是通过计算伯努利分布随机变量的序列而生成的。在一阶整数值自回归模型中, 算子的取值总是小于前一刻的取值, 所以在实际发生的数据中, 这就成为了二项稀疏算子的一个局限。因此, 有人提出用几何随机变量来表示这些事件。于是, Aly 和 Bouzar 和 Ristic 在 2009 年引入了负二项稀疏算子[4], 之后 Ristic 提出了算子服从负二项分布的一阶整数值自回归模型[5]。上述整数值时间序列模型的一个共同特点是它们都是单一算子模型。因此不能很好的拟合显示生活中的整数值时间序列数据。一些随机事件可能会以一定的概率在一个观察周期内滞留或者消失, 或者以一定的概率它们又会变得非常活跃, 并在一段时间后引发更多随机事件。为解决此类问题, Nastic 和 Ristic 提出了二项稀疏算子与负二项稀疏算子以一定概率混合的整数值自回归模型[6], 此模型的稀疏参数是固定的。虽然这种假定给研究带来了很大的方便, 但它显然是不合理的。因为数据的变化是受多方面环境因素的影响, 而环境是随时间变化的, 所以模型的稀疏参数不应该是固定不变的, 而应该是随机的, 所以在 2007 年, Zheng H, Basawa IV, Datta S 等人提出将一阶整数值自回归模型的系数随机化[7]。2008 年, Zheng H 建立了一阶观测驱动的整数值自回归模型[8]。2017 年, Han Li, Kai Yan 和 Shishun Zhao 等人提出一阶整数值自回归模型系数的随机化[9], 2018 年, Meiju Yu, Dehui Wang 等人提出基于负二项算子的观测驱动随机系数 INAR(1)模型[10]。

本文在上述模型的基础上, 对比优劣, 建立了一阶泊松随机系数混合算子整数值自回归模型 (Po-RCMTINAR(1)), 并且研究了模型的统计性质, 对模型参数运用 Yule-Walker 方法进行估计, 最后通过一组实际数据集来探讨模型的应用。

2. 混合算子

$\{W_i\}_{i \in \mathbb{N}}$ 是一个独立同分布的随机变量序列, 它的分布为[11]:

$$W_i = \begin{cases} B_i, & \text{以概率 } p \\ G_i, & \text{以概率 } 1-p \end{cases}, p \in [0, 1] \quad (1)$$

其中 B_i 和 G_i 分别表示服从 Bernoulli(α) 分布的随机变量和服从 $\text{geometric}\left(\frac{\alpha}{1+\alpha}\right)$ 分布的随机变量, G_i 的概率质量函数为 $P(G_i = x) = \frac{\alpha^x}{(1+\alpha)^{x+1}}$, $x = 0, 1, \dots, \alpha \in [0, 1]$, 假设 B_i 和 G_i 是独立随机变量。从(1)式中我们可以得出 W_i 是 Bernoulli 和 geometric 分布的混合, 当 $p = 0$ 时, 它服从 geometric 分布; 当 $p = 1$ 时, 它服从 Bernoulli 分布。

由形式(1)我们可以写出 W_i 的概率母函数:

$$\varphi_W(s) \equiv E(s^{W_i}) = p(1 - \alpha + \alpha s) + \frac{1 + p}{1 + \alpha - \alpha s} = \frac{1 - \alpha^2 p(1 - s)^2}{1 + \alpha(1 - s)}$$

现在我们给出混合算子定义[11]:

定义 2.1 运算符 $\alpha \cdot_p$ 被定义为

$$\alpha \cdot_p X = \sum_{i=1}^X W_i$$

其中 $\alpha \in [0, 1]$, X 是非负整数值随机变量且与 $\{W_i\}_{i \in \mathbb{N}}$ 独立。

3. 一阶泊松随机系数混合算子整数值自回归时间序列模型

一阶泊松随机系数混合算子整数值自回归模型(Po-RCMTINAR(1))被定义为:

$$X_t = \alpha_t \cdot_p X_{t-1} + \varepsilon_t, \quad t \geq 2 \quad (2)$$

这里 $\alpha_t \cdot_p X_{t-1}$ 是由定义 2.1 给出的混合算子; $\{\alpha_t\}$ 是 $[0, 1]$ 上独立同分布的随机序列具有概率分布 $\{P_\alpha\}$; $\{\varepsilon_t\}$ 是服从参数为 λ 的泊松分布随机变量序列; X_1 、 $\{\alpha_t\}$ 和 $\{\varepsilon_t\}$ 是相互独立的; 对于 $s > 0$, $\{\varepsilon_t\}$ 与 $\alpha_t \cdot_p X_{t-s}$ 是相互独立的。令 $E(\alpha_t) = \alpha$, $\text{Var}(\alpha_t) = \sigma_\alpha^2$, 并假设这些随机变量都是有限的。

下面研究模型中参数的统计性质。

模型的统计性质

$$1) E(X_t | X_{t-1}) = \alpha X_{t-1} + \lambda$$

$$2) E(X_t) = \frac{\lambda}{1 + \alpha}, \quad \text{如果 } E(X_1) = \frac{\lambda}{1 + \alpha}$$

$$3) \text{Var}(X_t | X_{t-1}, \alpha_t) = \alpha_t (1 - 2\alpha_t p + \alpha_t) X_{t-1} + \lambda$$

$$4) \text{Var}(X_t | X_{t-1}) = \alpha X_{t-1} - 2p(\sigma_\alpha^2 + \alpha^2) X_{t-1} + (\sigma_\alpha^2 + \alpha^2) X_{t-1}^2 + \lambda$$

$$5) \text{Var}(X_t) = \frac{b}{1 - (\sigma_\alpha^2 + \alpha^2)}, \quad \text{其中 } b = \frac{\alpha\lambda}{1 - \alpha} - \frac{(1 - 2p)(\sigma_\alpha^2 + \alpha^2)\lambda}{1 - \alpha} + \frac{\sigma_\alpha^2 \lambda^2}{1 - \alpha} + \lambda, \quad \text{如果}$$

$$\text{Var}(X_t) = \frac{b}{1 - (\sigma_\alpha^2 + \alpha^2)}$$

$$6) E(X_{t+k} | X_t) = \alpha^k X_t + \frac{\lambda(1 - \alpha^k)}{1 - \alpha}$$

$$7) \text{Cov}(X_{t+k}, X_t) = \alpha^k \text{Var}(X_t), \quad k \geq 0. \quad \text{如果 } \text{Var}(X_1) = \frac{b}{1 - (\sigma_\alpha^2 + \alpha^2)}, \quad \rho(k) = \alpha^k$$

证明:

$$\begin{aligned}
1) \quad E(X_t|X_{t-1}) &= E[\alpha_t \cdot_p X_{t-1} + \varepsilon_t | X_{t-1}] \\
&= E[E(\alpha_t \cdot_p X_{t-1} | X_{t-1}, \alpha_t) | X_{t-1}] + \lambda \\
&= E[\alpha_t X_{t-1}] + \lambda \\
&= \alpha X_{t-1} + \lambda
\end{aligned}$$

$$\begin{aligned}
2) \quad E(X_t) &= E[\alpha_t \cdot_p X_{t-1} + \varepsilon_t] \\
&= E[E(\alpha_t \cdot_p X_{t-1} | X_{t-1})] + \lambda \\
&= E[E(\alpha_t \cdot_p X_{t-1} | X_{t-1}, \alpha_t) | X_{t-1}] + \lambda \\
&= E(\alpha_t X_{t-1}) + \lambda \\
&= \alpha X_{t-1} + \lambda
\end{aligned}$$

$$3) \quad \text{Var}(X_t | X_{t-1}, \alpha_t) = E(X_t^2 | X_{t-1}, \alpha_t) + [E(X_t | X_{t-1}, \alpha_t)]^2$$

其中,

$$\begin{aligned}
E(X_t^2 | X_{t-1}, \alpha_t) &= E[(\alpha_t \cdot_p X_{t-1})^2 | X_{t-1}, \alpha_t] + 2E(\varepsilon_t \cdot \alpha_t \cdot_p X_{t-1} | X_{t-1}, \alpha_t) + E(\varepsilon_t^2 | X_{t-1}, \alpha_t) \\
&= \text{Var}(\alpha_t \cdot_p X_{t-1} | X_{t-1}, \alpha_t) + E(\alpha_t \cdot_p X_{t-1} | X_{t-1}, \alpha_t)^2 + 2\lambda \alpha_t X_{t-1} + \lambda + \lambda^2 \\
&= \alpha_t (1 - 2p\alpha_t + \alpha_t) X_{t-1} + \alpha_t^2 X_{t-1}^2 + 2\lambda \alpha_t X_{t-1} + \lambda + \lambda^2
\end{aligned}$$

所以,

$$\text{Var}(X_t | X_{t-1}, \alpha_t) = \alpha_t (1 - 2\alpha_t p + \alpha_t) X_{t-1} + \lambda$$

$$1) \quad \text{Var}(X_t | X_{t-1}) = E(X_t^2 | X_{t-1}) + [E(X_t | X_{t-1})]^2$$

其中

$$\begin{aligned}
E(X_t^2 | X_{t-1}) &= E[E(X_t^2 | X_{t-1}, \alpha_t) | X_{t-1}] \\
&= E[\alpha_t (1 - 2p\alpha_t + \alpha_t) X_{t-1} + \alpha_t^2 X_{t-1}^2 + 2\lambda \alpha_t X_{t-1} + \lambda + \lambda^2 | X_{t-1}] \\
&= \alpha X_{t-1} - 2p(\sigma_\alpha^2 + \alpha^2) X_{t-1} + (\sigma_\alpha^2 + \alpha^2) X_{t-1} + (\sigma_\alpha^2 + \alpha^2) X_{t-1}^2 + 2\lambda \alpha X_{t-1} + \lambda + \lambda^2
\end{aligned}$$

所以

$$\begin{aligned}
\text{Var}(X_t | X_{t-1}) &= \alpha X_{t-1} - 2p(\sigma_\alpha^2 + \alpha^2) X_{t-1} + (\sigma_\alpha^2 + \alpha^2) X_{t-1} + \sigma_\alpha^2 X_{t-1}^2 + \alpha^2 X_{t-1}^2 + 2\lambda \alpha X_{t-1} + \lambda + \lambda^2 \\
&= \alpha X_{t-1} - 2p(\sigma_\alpha^2 + \alpha^2) X_{t-1} + (\sigma_\alpha^2 + \alpha^2) X_{t-1} + \sigma_\alpha^2 X_{t-1}^2 + \lambda
\end{aligned}$$

$$\begin{aligned}
2) \quad \text{Var}(X_t) &= \text{Var}[E(X_t | X_{t-1})] + E[\text{Var}(X_t | X_{t-1})] \\
&= \alpha^2 \text{Var}(X_{t-1}) + \alpha E(X_{t-1}) - (1 - 2p)(\sigma_\alpha^2 + \alpha^2) E(X_{t-1}) + \sigma_\alpha^2 E(X_{t-1}^2) + \lambda \\
&= (\sigma_\alpha^2 + \alpha^2) \text{Var}(X_{t-1}) + \frac{\alpha \lambda}{1 - \alpha} - \frac{(1 - 2p)(\sigma_\alpha^2 + \alpha^2) \lambda}{1 - \alpha} + \frac{\sigma_\alpha^2 \lambda^2}{(1 - \alpha)^2} + \lambda \\
&= (\sigma_\alpha^2 + \alpha^2) \text{Var}(X_{t-2}) + (\sigma_\alpha^2 + \alpha^2) b + b \\
&= (\sigma_\alpha^2 + \alpha^2) \text{Var}(X_1) + \frac{b[1 - (\sigma_\alpha^2 + \alpha^2)^{t-1}]}{1 - (\sigma_\alpha^2 + \alpha^2)}
\end{aligned}$$

$$\text{当 } t \rightarrow \infty \text{ 时, } \text{Var}(X_t) = \frac{b}{1 - (\sigma_\alpha^2 + \alpha^2)},$$

$$\text{其中 } b = \frac{\alpha\lambda}{1-\alpha} - \frac{(1-2p)(\sigma_\alpha^2 + \alpha^2)\lambda}{1-\alpha} + \frac{\sigma_\alpha^2\lambda^2}{(1-\alpha)^2} + \lambda.$$

$$\begin{aligned} 3) \quad E(X_{t+k}|X_t) &= E(E(X_{t+k}|X_{t+k-1})|X_t) \\ &= E(\alpha X_{t+k-1} + \lambda | X_t) \\ &= \alpha E(X_{t+k-1}|X_t) + \lambda \\ &= \alpha^2 E(X_{t+k-1}|X_t) + \alpha\lambda + \lambda \\ &= \alpha^k X_t + \frac{\lambda(1-\alpha^k)}{1-\alpha} \end{aligned}$$

$$\begin{aligned} 4) \quad \text{Cov}(X_{t+k}, X_t) &= E(X_{t+k} X_t) - E(X_t) E(X_{t+k}) \\ &= E(E(X_{t+k} X_t | X_t)) - E(X_t) E(X_{t+k}) \\ &= E\left(X_t \left(\alpha^k X_t + \frac{\lambda(1-\alpha^k)}{1-\alpha}\right)\right) - E(X_t) \left[\alpha^k E(X_t) + \frac{\lambda(1-\alpha^k)}{1-\alpha}\right] \\ &= \alpha^k E(X_t^2) - \alpha^k [E(X_t)]^2 \\ &= \alpha^k \text{Var}(X_t) \end{aligned}$$

如果 $i=0$, $\alpha_{t+k} \cdot_p \cdots \alpha_{t+k-i+1} \cdot_p \varepsilon_{t+k-i} = \varepsilon_{t+k}$,

因此,

$$\begin{aligned} \text{Cov}(X_{t+k}, X_t) &= \text{Cov}\left(\alpha_{t+k} \cdot_p \cdots \alpha_{t+1} \cdot_p X_t + \sum_{i=0}^{k-1} \alpha_{t+k} \cdot_p \cdots \alpha_{t+k-i+1} \cdot_p \varepsilon_{t+k-i+1}, X_t\right) \\ &= E\left(X_t (\alpha_{t+k} \cdot_p \cdots \alpha_{t+1} \cdot_p X_t)\right) - E(X_t) E(\alpha_{t+k} \cdot_p \cdots \alpha_{t+1} \cdot_p X_t) \\ &= \alpha^k E(X_t^2) - \alpha^k E(X_t)^2 \\ &= \alpha^k \text{Var}(X_t) \end{aligned}$$

根据以上统计性质, 对模型中参数进行矩估计。

4. 参数估计

Yule-Walker 估计

对于模型(3.1)中的未知参数 $(\alpha, \lambda, p, \sigma_\alpha^2)$ 我们使用 Yule-Walker 去估计。由(3.1)中的模型性质可以得到 $\text{Corr}(X_t, X_{t-1}) = \alpha$ 和 $E(X_t) = \alpha E(X_{t-1}) + \lambda$, 可以直接得出 α 和 λ 的估计:

$$\begin{aligned} \hat{\alpha} &= \frac{\sum_{t=2}^n (X_t - \bar{X}_n)(X_{t-1} - \bar{X}_n)}{\sum_{t=1}^n (X_t - \bar{X}_n)^2}, \\ \hat{\lambda} &= \bar{X}_n - \hat{\alpha} \bar{X}_{n-1}. \end{aligned}$$

其中, $\bar{X}_n = \frac{1}{n} \sum_{t=1}^n X_t$, $\bar{X}_{n-1} = \frac{1}{n-1} \sum_{t=2}^n X_{t-1}$ 。另外, 为了估计, 我们需要进行一些计算得到下面的辅助公式:

$$\begin{aligned}
 E\left((\alpha_t \cdot_p X_{t-1})^2\right) &= (\sigma_\alpha^2 + \alpha^2)E(X_{t-1}^2) + [\alpha + (1-2p)(\sigma_\alpha^2 + \alpha^2)]E(X_{t-1}), \\
 E(X_{t-1}(\alpha_t \cdot_p X_{t-1})) &= \alpha E(X_{t-1}^2), \\
 E\left(X_{t-1}(\alpha_t \cdot_p X_{t-1})^2\right) &= (\sigma_\alpha^2 + \alpha^2)E(X_{t-1}^3) + [\alpha + (1-2p)(\sigma_\alpha^2 + \alpha^2)]E(X_{t-1}^2), \\
 E(X_{t-1}^2(\alpha_t \cdot_p X_{t-1})) &= \alpha E(X_{t-1}^3), \\
 E\left(X_{t-1}^2(\alpha_t \cdot_p X_{t-1})^2\right) &= (\sigma_\alpha^2 + \alpha^2)E(X_{t-1}^4) + [\alpha + (1-2p)(\sigma_\alpha^2 + \alpha^2)]E(X_{t-1}^3).
 \end{aligned}$$

通过上面这些辅助公式可得:

$$\text{Cov}(X_t^2, X_{t-1}) = (\sigma_\alpha^2 + \alpha^2)\text{Cov}(X_{t-1}^2, X_t^2) + [\alpha + (1-2p)(\sigma_\alpha^2 + \alpha^2) + 2\alpha\lambda] \text{Var}(X_{t-1}) \tag{①}$$

$$\text{Cov}(X_t^2, X_{t-1}^2) = (\sigma_\alpha^2 + \alpha^2)\text{Var}(X_{t-1}^2) + [\alpha + (1-2p)(\sigma_\alpha^2 + \alpha^2) + 2\alpha\lambda] \text{Cov}(X_{t-1}^2, X_{t-1}) \tag{②}$$

令上述①式乘以 $\text{Cov}(X_{t-1}^2, X_{t-1})$ 减去②式乘以 $\text{Var}(X_{t-1})$, 解出:

$$\begin{aligned}
 \hat{\sigma}_\alpha^2 &= \frac{\text{C}\hat{\text{ov}}(X_t^2, X_{t-1}^2)\text{V}\hat{\text{ar}}(X_{t-1}) - \text{C}\hat{\text{ov}}(X_t^2, X_{t-1})\text{C}\hat{\text{ov}}(X_{t-1}^2, X_{t-1})}{\text{V}\hat{\text{ar}}(X_{t-1}^2)\text{V}\hat{\text{ar}}(X_{t-1}) - [\text{C}\hat{\text{ov}}(X_{t-1}^2, X_{t-1})]^2} - \hat{\alpha}^2, \\
 \hat{p} &= \frac{(\hat{\sigma}_\alpha^2 + \hat{\alpha}^2)\text{C}\hat{\text{ov}}(X_{t-1}^2, X_{t-1}) + (\hat{\alpha} + 2\hat{\alpha}\hat{\lambda} + \hat{\sigma}_\alpha^2 + \hat{\alpha}^2)\text{V}\hat{\text{ar}}(X_{t-1}) - \text{C}\hat{\text{ov}}(X_t^2, X_{t-1})}{2(\hat{\sigma}_\alpha^2 + \hat{\alpha}^2)\text{V}\hat{\text{ar}}(X_{t-1})}.
 \end{aligned}$$

其中,

$$\begin{aligned}
 \text{C}\hat{\text{ov}}(X_t^2, X_{t-1}^2) &= \frac{1}{n-1} \sum_{t=2}^n X_t^2 X_{t-1}^2 - \left(\frac{1}{n-1} \sum_{t=2}^n X_t^2 \right) \left(\frac{1}{n-1} \sum_{t=2}^n X_{t-1}^2 \right), \\
 \text{V}\hat{\text{ar}}(X_{t-1}) &= \frac{1}{n-2} \sum_{t=2}^n (X_{t-1} - \bar{X}_{n-1})^2, \\
 \text{C}\hat{\text{ov}}(X_t^2, X_{t-1}) &= \frac{1}{n-1} \sum_{t=2}^n X_t^2 X_{t-1} - \left(\frac{1}{n-1} \sum_{t=2}^n X_t^2 \right) \left(\frac{1}{n-1} \sum_{t=2}^n X_{t-1} \right), \\
 \text{C}\hat{\text{ov}}(X_{t-1}^2, X_{t-1}) &= \frac{1}{n-1} \sum_{t=2}^n X_{t-1}^3 - \left(\frac{1}{n-1} \sum_{t=2}^n X_{t-1}^2 \right) \left(\frac{1}{n-1} \sum_{t=2}^n X_{t-1} \right), \\
 \text{V}\hat{\text{ar}}(X_{t-1}^2) &= \frac{1}{n-2} \sum_{t=2}^n (X_{t-1}^2 - \bar{X}_{n-2})^2, \\
 \bar{X}_{n-2} &= \frac{\sum_{t=2}^n X_{t-1}^2}{n-1}.
 \end{aligned}$$

5. 案例分析

在这一部分, 基于一组实际数据集来分析提出的一阶泊松随机系数混合算子整数自回归模型 (Po-RCMTINAR(1)) 的实用性。该数据来自预测原理网站 (<http://www.forecastingprinciples.com>)。并且我们用 MSE 准则评价模型, 其 MSE 定义为 [12]:

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

这里, $\hat{X}_t = E(X_k | \mathcal{F}_{t-1})$ 并且取 $m = 30$, $\mathcal{F}_t = \sigma(X_1, X_2, \dots, X_t)$ 。

此组数据来自犯罪数据 PghTracts 中 CMIS 的第 97 组数据, 是从 1990 年 1 月到 2001 年 12 月, 共有 144 个观测数据。图 1 给出了该数据的路径图, 图 2 是其自相关图和偏自相关图。根据偏自相关函数在滞后 2 阶开始平稳, 说明这组数据适用于一阶模型对其进行分析。我们使用 Po-RCMTINAR(1)模型与一阶泊松混合算子模型(Po-MTINAR(1))进行分析此数据, 得到相应参数估计值和 MSE 准则值如表 1。

在图 3 中, 给出了两种模型残差的自相关图和偏自相关图。在这里, 我们用到标准化皮尔逊残差的公式[13]:

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

如图 3 所示, 残差序列是平稳序列。此外, Po-RCMTINAR(1)模型和 Po-MTINAR(1)模型的期望和方差分别为 $(-0.0103, 1.3513)$ 和 $(-0.0032, 0.2587)$, 说明模型 Po-RCMTINAR(1)更接近正态分布。由表 1 的结果知 Po-RCMTINAR(1)模型的 MSE 准则值比 Po-MTINAR(1)模型小, 所以一阶泊松随机系数混合算子整值自回归模型(Po-RCMTINAR(1))更适合被用于分析此数据。

Table 1. Value of parameters' estimation and MSE
表 1. 参数估计值及 MSE 值

模型	Yule-Walker 估计	MSE
Po-RNMTINAR(1)	$\hat{\alpha} = 0.2703$	7.1507
	$\hat{\sigma}_\alpha^2 = 0.0298$	
	$\hat{\lambda} = 4.4110$	
	$\hat{\rho} = 0.3743$	
Po-MTINAR(1)	$\hat{\alpha} = 0.2703$	7.1529
	$\hat{\lambda} = 6.0348$	
	$\hat{\rho} = 5.2578$	

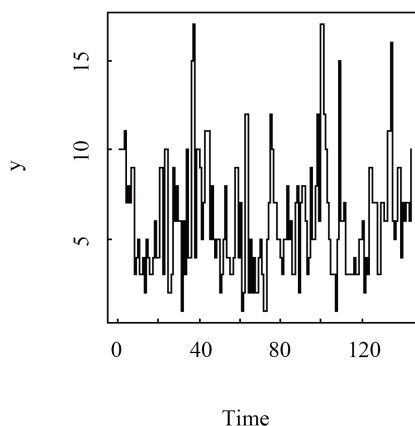


Figure 1. Plot of counts

图 1. 样本图

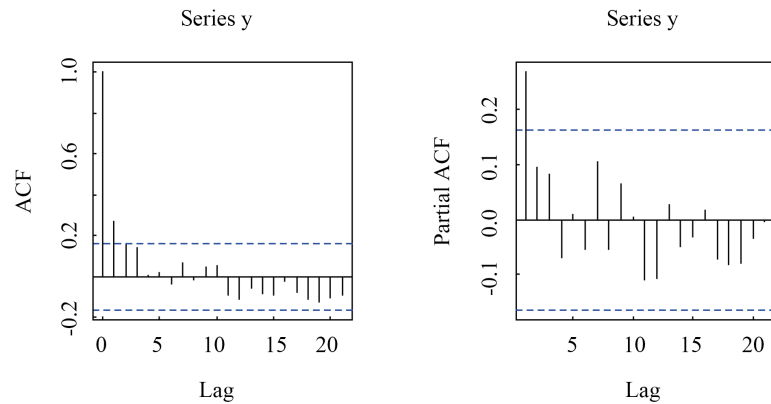


Figure 2. ACF and PACF plots of counts
图 2. 数据 ACF 图和 PACF 图

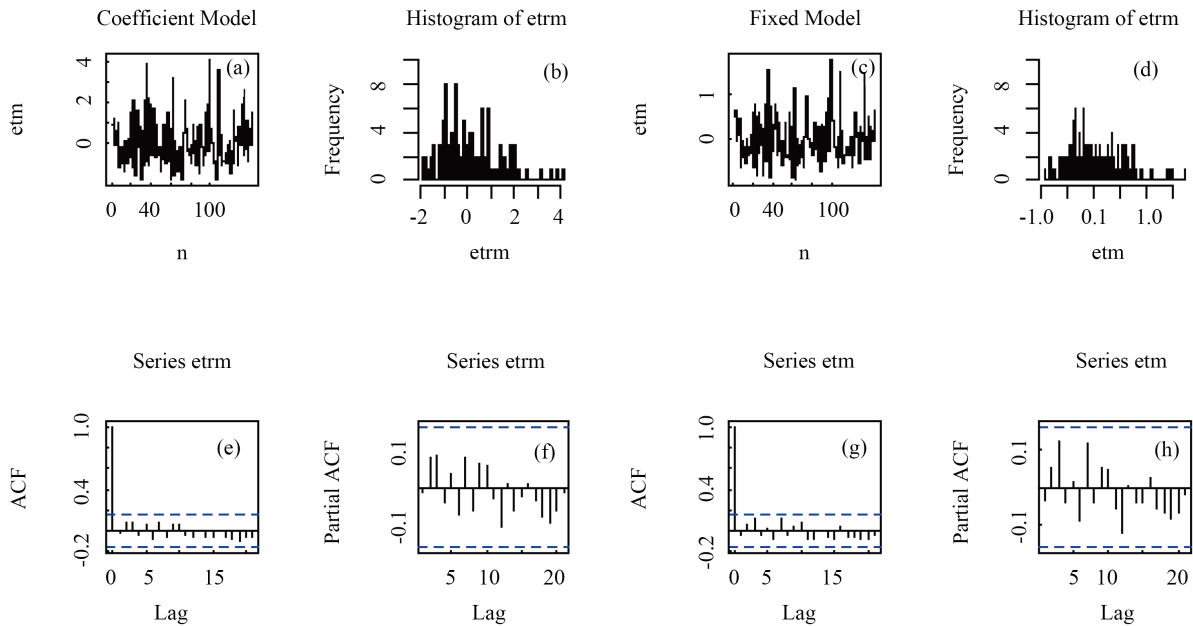


Figure 3. (a), (c) Standardized residuals; (b), (d) Histograms of standardized residuals; (e), (g) ACF plots of residuals; (f), (h) PACF plots of residuals

图 3. (a)、(c)为标准化残差图；(b)、(d)为标准化残差直方图；(e)、(g)为残差的自相关图；(f)、(h)为残差的偏自相关图

6. 结论

在本文中，我们对以往的整数值自回归模型进行优化：提出了一阶泊松随机系数混合算子整数值自回归模型，其自回归系数是随时间变化的。并且推导和证明了该模型的统计性质，利用 Yule-Walker 方法对模型的参数进行了估计。最后，我们将一阶泊松随机系数混合算子整数值自回归模型与一阶泊松混合算子整数值自回归模型用于分析真实数据进行比较。结果表明，一阶随机系数混合算子整数值自回归模型 (Po-RCMTINAR(1)) 更适合分析此组真实数据。

致 谢

衷心地感谢我的研究生导师赵丽华老师和给我很多帮助的刘秀芳老师，感谢她们在整个学习及论文

写作过程中, 对我认真的指导, 教会我很多科研知识; 在生活中, 经常告诫我不要有太大压力, 做好自己, 让我在漫漫学涯中感受到了温暖和肯定。其次, 感谢给予转载和引用权的文献及研究思想和设想的所有作者, 正是借助你们的肩膀, 我才能够更好地完成论文的撰写。

参考文献

- [1] 于梅菊. 随机系数整值自回归过程的推广及其统计推断[D]: [博士学位论文]. 吉林: 吉林大学数学系, 2019.
- [2] Steutel, F.W. and Van Harn, K. (1979) Discrete Analogues of Self-Decomposability and Stability. *Annals of Probability*, **7**, 893-899. <https://doi.org/10.1214/aop/1176994950>
- [3] 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. <https://doi.org/10.1111/j.1467-9892.1987.tb00438.x>
- [4] Aly, E.A. and Bouzar, N. (1994) Explicit Stationary Distributions for Some Galton-Watson Processes with Immigration. *Stochastic Models*, **10**, 499-517. <https://doi.org/10.1080/15326349408807305>
- [5] Ristic, M.M., Bakouch, H.S., Nastic, A.S., *et al.* (2009) A New Geometric First-Order Integer-Valued Autoregressive (NGINAR(1)) Process. *Journal of Statistical Planning and Inference*, **139**, 2218-2226. <https://doi.org/10.1016/j.jspi.2008.10.007>
- [6] Nastic, A.S., Ristic, M. and Janjic, A.D. (2017) A Mixed Thinning Based Geometric INAR(1) Model. *Filomat*, **31**, 4009-4022. <https://doi.org/10.2298/FIL1713009N>
- [7] 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. <https://doi.org/10.1016/j.jspi.2005.12.003>
- [8] Zheng, H. and Basawa, I.V. (2008) First-Order Observation-Driven Integer-Valued Autoregressive Processes. *Statistics and Probability Letters*, **78**, 1-9. <https://doi.org/10.1016/j.spl.2007.04.017>
- [9] Li, H., Yang, K., Zhao, S.S. and Wang, D.H. (2017) First-Order Random Coefficients Integer-Valued Threshold Autoregressive Processes. *AStA Advances in Statistical Analysis*, **102**, 305-331. <https://doi.org/10.1007/s10182-017-0306-3>
- [10] Yu, M.J., Wang, D.H. and Yang, K. (2018) A Class of Observation-Driven Random Coefficient INAR(1) Processes Based on Negative Binomial Thinning. *Journal of the Korean Statistical Society*, **48**, 248-264. <https://doi.org/10.1016/j.jkss.2018.11.004>
- [11] Li, C., Wang, D., Zhang, H., *et al.* (2015) First-Order Mixed Integer-Valued Autoregressive Processes with Zero-Inflated Generalized Power Series Innovations. *Journal of the Korean Statistical Society*, **44**, 232-246. <https://doi.org/10.1016/j.jkss.2014.08.004>
- [12] Liu, X.F. and Wang, D.H. (2019) Estimation of Parameters in the DDRCINAR(p) Model. *Brazilian Journal of Probability and Statistics*, **33**, 638-673. <https://doi.org/10.1214/18-BJPS405>
- [13] Liu, X., Wang, D., Deng, D., *et al.* (2020) Maximum Likelihood Estimation of the DDRCINAR(p) Model. *Communications in Statistics—Theory and Methods*, 1-25. <https://doi.org/10.1080/03610926.2020.1741627>