



A new estimation for INAR(1) process with Poisson distribution

Feilong Lu^{1,3} · Dehui Wang²

Received: 3 February 2021 / Accepted: 15 September 2021 / Published online: 28 September 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

The first-order Poisson autoregressive model may be suitable in situations where the time series data are non-negative integer valued. In this article, we propose a new parameter estimator based on empirical likelihood. Our results show that it can lead to efficient estimators by making effective use of auxiliary information. As a by-product, a test statistic is given, testing the randomness of the parameter. The simulation values show that the proposed test statistic works well. We have applied the suggested method to a real count series.

Keywords Integer-valued time series · Empirical likelihood · Test statistic

1 Introduction

Owen (1988) first proposed an empirical likelihood (EL) as a way to extend likelihood-based inference ideas to certain nonparametric situations. The advantages of the empirical likelihood are now recognized. For more details and further studies, we refer to the book or the papers by Owen (1991), Qin and Lawless (1994), Qin and Lawless (1995), and Zhang et al. (2011b) among others.

In many different fields, for example, economics, medicine, actuarial statistics etc, lots of interesting variables are integer-valued. Recently, there has been an increasing attention in modeling non-negative integer-valued time series. Some models have been proposed in the literature. See, for instance, (Weiss 2017) and (Steutel and Van Harn

✉ Dehui Wang
wangdehui@lnu.edu.cn; wangdh@jlu.edu.cn

¹ School of Science, University of Science and Technology Liaoning, Anshan 114051, China

² School of Economics, Liaoning University, Shenyang 110036, China

³ School of Mathematics, Jilin University, Changchun 130012, China

1979) proposed a thinning operator “ \circ ”, which is defined as

$$\phi \circ X = \sum_{i=1}^X B_i,$$

where $\phi \in [0, 1)$, X is an integer-valued random variable and $\{B_i\}$ is an independent and identically distributed Bernoulli random sequence with $P(B_i = 1) = 1 - P(B_i = 0) = \phi$ and is independent of X . With the operator, the first-order integer-valued autoregressive model (INAR(1)) is defined as follows

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

where the innovations $\{Z_t\}$ is a sequence of i.i.d. non-negative integer-valued random variables with mean λ and variance σ_Z^2 . Here all thinning operations are performed independently of each other and of $\{Z_t\}$ and the thinning operations at each time t as well as Z_t are independent of $\{X_s\}_{s < t}$. The model has some similar properties as the ordinary AR(1) model and it has been discussed by Al-Osh and ALZAID (1987), Alzaid and Al-Osh (1988), and Al-Osh and Aly (1992). If the $\{Z_t\}$ are assumed to be Poisson distributed, $Z_t \sim Poi(\lambda)$ with $\lambda > 0$, then $\{X_t\}$ is a stationary and ergodic Markov chain with a Poisson marginal distribution, such that mean and variance are equal to each other, $\mu_X = \sigma_X^2 = \lambda / (1 - \phi)$, and all moments exist. For conditional mean and variance, we have (Alzaid and Al-Osh (1988)): $E[X_t | X_{t-1}] = \phi \cdot X_{t-1} + \lambda$, $\text{Var}[X_t | X_{t-1}] = \phi(1 - \phi) \cdot X_{t-1} + \lambda$.

In this article, based on empirical likelihood, we provide a novel estimator which has small bias and is efficient by making use of auxiliary information shown in the following section. (Keith Freeland and McCabe 2005) derived the asymptotic properties of the conditional least squares (CLS) estimators in this model. (Bourguignon and Vasconcellos 2015) proposed a bias-adjusted estimator.

Note that the parameter ϕ may vary with time and it may be random. Zheng et al. (2007) proposed the following first-order random coefficient integer-valued autoregressive (RCINAR(1)) model:

$$X_t = \phi_t \circ X_{t-1} + Z_t, \quad t \geq 1, \quad (2)$$

where $\{\phi_t\}$ is an i.i.d. sequence with cumulative distribution function on $[0, 1)$, with $E[\phi_t] = \phi$ and $\text{Var}(\phi_t) = \sigma_\phi^2$; $\{Z_t\}$ is an i.i.d. non-negative integer-valued sequence as the same in the model (1). Moreover, $\{\phi_t\}$ and $\{Z_t\}$ are independent. The process $\{X_t\}$ is also a stationary and ergodic Markov chain. For conditional mean and variance, we have $E[X_t | X_{t-1}] = \phi \cdot X_{t-1} + \lambda$, $\text{Var}[X_t | X_{t-1}] = \sigma_\phi^2 \cdot X_{t-1}^2 + (\phi(1 - \phi) - \sigma_\phi^2) \cdot X_{t-1} + \sigma_Z^2$. In Zheng et al. (2007), the authors consider two methods, conditional least squares and the modified quasi-likelihood (MQL). Zhang et al. (2011a) studied the empirical likelihood method for the model (2). Kang and Lee (2009) considered the problem of testing for a parameter change in the model (2). There is one of the problems of interest, given the data $\{x_1, x_2, \dots, x_n\}$, test the hypothesis that there is no time variation for the thinning parameter ϕ (the series is INAR(1)) against that it

varies randomly across the time (it is RCINAR(1)). This is essentially same as testing $H_0 : \sigma_\phi^2 = 0$ against $H_1 : \sigma_\phi^2 > 0$.

In this article, we give a test statistic for testing this hypothesis. Zhao and Hu (2015) have proposed a test for randomness of the coefficient of an RCINAR(1) process, but using the least squares estimates of the parameter σ_ϕ^2 . Awale et al. (2019) developed a locally most powerful-type testing the hypothesis that the thinning parameter is constant across the time depended on the distribution of the innovation Z_t .

The rest of the paper is organized as follows. In Sect. 2, we introduce the methodology and the main results. In Sect. 3, simulation results are reported. In Sect. 4, an empirical example is discussed.

Throughout the paper, we use the notation “ \xrightarrow{d} ” to denote convergence in distribution, “ $'$ ” to denote the transpose operator and $\| \cdot \|$ to denote the Euclidean norm of the matrix.

2 Methodology and main results

Al-Osh and ALZAID (1987) defined the Poisson AR(1) model by

$$X_t = \phi \circ X_{t-1} + Z_t, \quad t \geq 1, \tag{3}$$

where $\phi \circ X_{t-1} = \sum_{i=1}^{X_{t-1}} B_{i,t}$, $B_{1,t}, B_{2,t}, \dots, B_{X_{t-1},t}$ are i.i.d. Bernoulli random variables with $P(B_{it} = 1) = 1 - P(B_{it} = 0) = \phi$; $\{Z_t\}$ are i.i.d. with a Poisson distribution having mean λ and are independent of the Bernoulli variables B_{it} .

Remark 1 $\{X_t\}$ is a Markov chain on $\{0, 1, 2, \dots\}$ with the transition probability,

$$p_{ij} = P(X_t = j | X_{t-1} = i) = \sum_{k=0}^{\min\{i,j\}} C_i^k \phi^k (1 - \phi)^{i-k} \frac{\lambda^{(j-k)} e^{-\lambda}}{(j - k)!}.$$

For PINAR(1) process (3), let $\theta = [\phi, \lambda]'$, we can construct estimation equations:

$$\mathbf{m}_t(\theta) = \begin{pmatrix} m_{1t}(\theta) \\ m_{2t}(\theta) \\ m_{3t}(\theta) \end{pmatrix} = \mathbf{0},$$

where $m_{1t}(\theta) = X_t - \phi X_{t-1} - \lambda$, $m_{2t}(\theta) = (X_t - \phi X_{t-1} - \lambda) X_{t-1}$, $m_{3t}(\theta) = (X_t - \phi X_{t-1} - \lambda)^2 - \phi(1 - \phi) X_{t-1} - \lambda$.

The profile empirical likelihood ratio function is (Owen 1991)

$$R(\theta) = \max \left\{ \prod_{t=1}^n n p_t \mid p_t \geq 0, \sum_{t=1}^n p_t = 1, \sum_{t=1}^n p_t m_t(\theta) = 0 \right\}. \tag{4}$$

The maximum may be obtained via Lagrange multipliers. Let

$$H = \sum_{t=1}^n \log(np_t) - n\boldsymbol{\beta}' \left(\sum_{t=1}^n p_t m_t(\boldsymbol{\theta}) \right) + \gamma \left(1 - \sum_{t=1}^n p_t \right),$$

where γ and $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3)'$ are Lagrange multipliers. Taking the partial derivative of H with respect to p_t to be zero, we have $\frac{\partial H}{\partial p_t} = \frac{1}{p_t} - n\boldsymbol{\beta}' m_t(\boldsymbol{\theta}) - \gamma = 0$, so

$$0 = \sum_{t=1}^n p_t \frac{\partial H}{\partial p_t} = n - \gamma, \quad p_t = \frac{1}{n} \frac{1}{1 + \boldsymbol{\beta}' m_t(\boldsymbol{\theta})}, \quad t = 1, 2, \dots, n.$$

Here, being a function of $\boldsymbol{\theta}$, $\boldsymbol{\beta} = \boldsymbol{\beta}(\boldsymbol{\theta})$ is with the restriction that

$$\frac{1}{n} \sum_{t=1}^n \frac{m_t(\boldsymbol{\theta})}{1 + \boldsymbol{\beta}' m_t(\boldsymbol{\theta})} = 0.$$

The empirical log-likelihood ratio (ELR) is

$$\ell(\boldsymbol{\theta}) = \log R(\boldsymbol{\theta}) = \sum_{t=1}^n \log(1 + \boldsymbol{\beta}' m_t(\boldsymbol{\theta})). \quad (5)$$

The asymptotic distribution of the ELR was derived by Qin and Lawless (1994) for general estimating equations and by Zhang et al. (2011b) for RCINAR(p) model. Basing on their results, it is possible to prove an asymptotic chi-squared distribution for (5) as shown in the following theorem.

Theorem 1 *As $n \rightarrow \infty$, we have $2\ell(\boldsymbol{\theta}_0) \xrightarrow{d} \chi^2(2)$, where $\ell(\boldsymbol{\theta})$ is defined in (5).*

So, it is easy to obtain the confidence region for the parameter $\boldsymbol{\theta}$. The result is as follows:

Corollary 1 *For the parameter $\boldsymbol{\theta}$, the $100(1 - \alpha)\%$ confidence region is $C_{\alpha,n} = \{\boldsymbol{\theta} | 2\ell(\boldsymbol{\theta}) \leq \chi_{1-\alpha}^2(2)\}$, where $1 - \alpha$ is the confidence level and $\chi_{1-\alpha}^2(2)$ is the upper $1 - \alpha$ -quantile of the chi-squared distributed with 2 degrees of the freedom.*

Then, we may find the maximum empirical likelihood estimator (MELE) $\check{\boldsymbol{\theta}}$ for the parameter $\boldsymbol{\theta}$ by minimizing $\ell(\boldsymbol{\theta})$ over $\boldsymbol{\theta}$.

Theorem 2 *For the MELE $\check{\boldsymbol{\theta}}$, we have*

$$\sqrt{n}(\check{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{d} N(\mathbf{0}, V),$$

where $V = [\boldsymbol{\Omega}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\Omega}]^{-1}$, $\boldsymbol{\Omega} = E \left(\frac{\partial \mathbf{m}_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)$, $\boldsymbol{\Sigma} = E [\mathbf{m}_t(\boldsymbol{\theta}) \mathbf{m}_t(\boldsymbol{\theta})']$.

Remark 2 If we don't consider the third equation and only use $m_{1t} = 0, m_{2t} = 0$ as the estimation equations, the maximum empirical likelihood estimator is consistent with the conditional least squares estimator.

Remark 3 The conditional least squares estimator (CLE) of the model's parameters can be calculated by the following expression

$$\hat{\phi} = \frac{n \sum_{t=1}^n X_t X_{t-1} - (\sum_{t=1}^n X_{t-1})(\sum_{t=1}^n X_t)}{n \sum_{t=1}^n X_{t-1}^2 - (\sum_{t=1}^n X_{t-1})^2}, \quad \hat{\lambda} = \frac{1}{n} \left(\sum_{t=1}^n X_t - \hat{\phi} \sum_{t=1}^n X_{t-1} \right).$$

The CLS estimator is consistent and asymptotically normal, see (Keith Freeland and McCabe 2005).

Remark 4 Note that the conditional variance is

$$V_{\theta}(X_t|X_{t-1}) := \text{Var}(X_t|X_{t-1}) = \phi(1 - \phi)X_{t-1} + \lambda,$$

then the maximum quasi-likelihood (MQL) estimator of the parameters are:

$$\begin{pmatrix} \tilde{\phi} \\ \tilde{\lambda} \end{pmatrix} = \left(\begin{matrix} \sum_{t=1}^n X_{t-1}^2 V_{\hat{\theta}}^{-1}(X_t|X_{t-1}) & \sum_{t=1}^n X_{t-1} V_{\hat{\theta}}^{-1}(X_t|X_{t-1}) \\ \sum_{t=1}^n X_{t-1} V_{\hat{\theta}}^{-1}(X_t|X_{t-1}) & \sum_{t=1}^n V_{\hat{\theta}}^{-1}(X_t|X_{t-1}) \end{matrix} \right)^{-1} \\ \times \begin{pmatrix} \sum_{t=1}^n X_{t-1} X_t V_{\hat{\theta}}^{-1}(X_t|X_{t-1}) \\ \sum_{t=1}^n X_t V_{\hat{\theta}}^{-1}(X_t|X_{t-1}) \end{pmatrix},$$

where $\hat{\theta}$ is its consistent estimator, in practice, we can use the conditional least squares estimator of θ . Similar to the proof of theorem 3.2 in Zheng et al. (2007), it can be proved that the estimator is also consistent and asymptotically normal.

Remark 5 The conditional maximum likelihood (CML) estimator of the parameter $\bar{\theta}$ is

$$\bar{\theta} = \arg \max \sum_{t=1}^n \log p_{x_t|x_{t-1}}(\theta),$$

where

$$p_{x_t|x_{t-1}}(\theta) := P(X_t = x_t|X_{t-1} = x_{t-1}) = \sum_{k=0}^{\min(x_t, x_{t-1})} \binom{x_{t-1}}{k} \phi^k (1 - \phi)^{x_{t-1}-k} \frac{\lambda^{x_t-k}}{(x_t - k)!}.$$

The CML estimator is consistent and asymptotically normal, see Freeland and McCabe (2004).

In order to test $H_0 : \sigma_{\phi}^2 = 0$, we may consider the following theorem.

Theorem 3 Under H_0 , we have $2\ell(\check{\theta}) \xrightarrow{d} \chi^2(1)$, as $n \rightarrow \infty$, where $\ell(\theta)$ is defined in (5).

Then, it follows that the null hypothesis $H_0 : \sigma_\phi^2 = 0$ may be tested using the statistic $T = 2\ell(\check{\theta})$ and the rejection region $\{T > \chi_{1-\alpha}^2(1)\}$.

3 Simulation study

Simulation studies are reported in the section to evaluate the finite sample performance of the proposed maximum empirical likelihood estimator (MELE). We computed the empirical bias and root of the mean squared errors (RMSE) based on 1000 replications for each parameter combination. These values are reported within parenthesis in Tables 1. For succinctness, we only report here partial results, to illustrate the effect of parameter ϕ while holding the other parameter λ identical. Other results are available from the authors upon request. In the following Table 1, the format (Bias, RMSE) is used; for example, (0.0142, 0.1204) means that the bias is 0.0142 and RMSE is 0.1204.

From the above results, we can see that MELE is a good estimation method producing estimators whose biases and RMSEs are overall comparable to the CML method. As expected, the CML performs best which has the smallest bias and RMSE and there are almost no differences between the MELE and CML. The CLS and QML have the same performance in terms of both bias and the RMSE. However, when the sample size is small or the parameters are large, the MELE works better than the CLS and QML.

In order to assess the performance of the suggested test statistic, we have done a lot of simulation studies. We simulate observations from the Poisson INAR(1) model various combinations of the parameters, as shown in 2. In each case, the sizes of samples are 100, 300, 500, 1000, 5000 and 1000 simulations were carried out. The distribution of ϕ_t is assumed to $Beta(a, b)$ under the alternative hypothesis. The empirical level and power of the test statistic is reported in Tables 2 and 3, respectively.

It can be seen that the test maintains the level and the power tends to one as sample size increase. We also perform the test proposed by Awale et al. (2019), since it is better than the test given by Zhao and Hu (2015). The results are reported in Tables 4 and 5.

From comparison with level and power of the the suggested test from Tables 2 and 3, it can be seen that the test is not as good as the proposed test using the probability distribution function, as accepted. However, the suggested test in this article is competitive and convenient as opposed to that of Awale et al. (2019).

4 Real data example

In this section, we employ the programme suggested in the previous section a real time series data. The data set consists of monthly counts of claimants collecting short-term wage loss benefit for born-related injuries received in workplace. The data set contains 120 observations for the period January 1985 to December 1994. These data were previously studied by Freeland and McCabe (2004) and Bourguignon and

Table 1 Estimates for parameter ϕ and $\lambda = 1$

	CLS	MQL	MELLE	CML
ϕ		$\phi = 0.1$		
$n = 50$	(-0.0235, 0.1416)	(-0.0205, 0.1459)	(0.0142, 0.1204)	(0.0112, 0.1143)
$n = 100$	(-0.0138, 0.0969)	(-0.0138, 0.0969)	(0.0013, 0.0856)	(-0.0011, 0.0836)
$n = 200$	(-0.0091, 0.0726)	(-0.0090, 0.0719)	(-0.0029, 0.0676)	(-0.0036, 0.0656)
$n = 500$	(-0.0017, 0.0451)	(-0.0018, 0.0449)	(-0.0008, 0.0450)	(-0.0012, 0.0441)
λ				
$n = 50$	(0.0252, 0.2112)	(0.0216, 0.2126)	(-0.0150, 0.1986)	(-0.0126, 0.1891)
$n = 100$	(0.0109, 0.1495)	(0.0110, 0.1501)	(-0.0117, 0.1434)	(-0.0028, 0.1403)
$n = 200$	(0.0072, 0.1054)	(0.0072, 0.1047)	(-0.0021, 0.1013)	(0.0012, 0.1003)
$n = 500$	(0.0041, 0.0671)	(0.0042, 0.0671)	(0.0024, 0.0672)	(0.0034, 0.0663)
		$\phi = 0.3$		
ϕ				
$n = 50$	(-0.0414, 0.1387)	(-0.0406, 0.1376)	(-0.0255, 0.1369)	(-0.0289, 0.1300)
$n = 100$	(-0.0164, 0.1030)	(-0.0157, 0.1017)	(-0.0096, 0.0997)	(-0.0097, 0.0945)
$n = 200$	(-0.0087, 0.0727)	(-0.0074, 0.0710)	(-0.0054, 0.0704)	(-0.0033, 0.0669)
$n = 500$	(-0.0040, 0.0453)	(-0.0030, 0.0446)	(-0.0021, 0.0419)	(-0.0012, 0.0404)
λ				
$n = 50$	(0.0589, 0.2436)	(0.0575, 0.2411)	(0.0205, 0.2446)	(0.0408, 0.2305)
$n = 100$	(0.0109, 0.1639)	(0.0177, 0.1636)	(0.0027, 0.1628)	(0.0095, 0.1570)
$n = 200$	(0.0072, 0.1213)	(0.0080, 0.1195)	(0.0036, 0.1177)	(0.0019, 0.1125)
$n = 500$	(0.0041, 0.0752)	(0.0039, 0.0744)	(0.0023, 0.0704)	(0.0017, 0.0686)

Table 1 continued

	CLS	MQL	MELLE	CML
ϕ				
$n = 50$	(-0.0512, 0.1367)	$\phi = 0.5$ (-0.0478, 0.1382)	(-0.0155, 0.1219)	(-0.0172, 0.1133)
$n = 100$	(-0.0267, 0.0970)	(-0.0256, 0.0950)	(-0.0092, 0.0830)	(-0.0113, 0.0777)
$n = 200$	(-0.0163, 0.0627)	(-0.0148, 0.0615)	(-0.0071, 0.0556)	(-0.0074, 0.0527)
$n = 500$	(-0.0029, 0.0419)	(-0.0026, 0.0404)	(-0.0001, 0.0346)	(-0.0001, 0.0334)
λ				
$n = 50$	(0.0882, 0.2997)	(0.0817, 0.3022)	(-0.0014, 0.2568)	(0.0195, 0.2451)
$n = 100$	(0.0456, 0.2074)	(0.0437, 0.2046)	(0.0070, 0.1771)	(0.0149, 0.1685)
$n = 200$	(0.0256, 0.1339)	(0.0227, 0.1324)	(0.0055, 0.1219)	(0.0079, 0.1152)
$n = 500$	(0.0055, 0.0909)	(0.0050, 0.0881)	(-0.0005, 0.0759)	(-0.0002, 0.0735)
ϕ		$\phi = 0.7$		
$n = 50$	(-0.0686, 0.1145)	(-0.0657, 0.1134)	(-0.0089, 0.0787)	(-0.0111, 0.0719)
$n = 100$	(-0.0330, 0.0756)	(-0.0296, 0.0747)	(-0.0050, 0.0495)	(-0.0059, 0.0474)
$n = 200$	(-0.0155, 0.0540)	(-0.0146, 0.0526)	(-0.0033, 0.0350)	(-0.0036, 0.0331)
$n = 500$	(-0.0054, 0.0337)	(-0.0053, 0.0325)	(-0.0004, 0.0223)	(-0.0007, 0.0215)
λ				
$n = 50$	(0.2214, 0.4130)	(0.2119, 0.4093)	(0.0104, 0.2816)	(0.0270, 0.2543)
$n = 100$	(0.0993, 0.2704)	(0.0882, 0.2690)	(0.0003, 0.1744)	(0.0079, 0.1667)
$n = 200$	(0.0424, 0.1853)	(0.0396, 0.1821)	(0.0008, 0.1215)	(0.0029, 0.1163)
$n = 500$	(0.0198, 0.1185)	(0.0192, 0.1137)	(0.0029, 0.0793)	(0.0035, 0.0752)

Table 1 continued

	CLS	MQL	MELE	CML
ϕ		$\phi = 0.9$		
$n = 50$	(-0.0850, 0.1002)	(-0.0846, 0.1000)	(-0.0013, 0.0568)	(-0.0047, 0.0249)
$n = 100$	(-0.0392, 0.0578)	(-0.0390, 0.0578)	(-0.0045, 0.0308)	(-0.0015, 0.0171)
$n = 200$	(-0.0203, 0.0352)	(-0.0119, 0.0349)	(-0.0016, 0.0137)	(-0.0013, 0.0119)
$n = 500$	(-0.0086, 0.0212)	(-0.0082, 0.0210)	(-0.0009, 0.0076)	(-0.0008, 0.0072)
λ				
$n = 50$	(0.8427, 1.0328)	(0.8376, 1.0298)	(0.1175, 0.5893)	(0.0368, 0.2521)
$n = 100$	(0.3885, 0.5873)	(0.3859, 0.5873)	(0.0375, 0.3121)	(0.0090, 0.1703)
$n = 200$	(0.2018, 0.3630)	(0.1970, 0.3594)	(0.0122, 0.1443)	(0.0096, 0.1192)
$n = 500$	(0.0832, 0.2170)	(0.0796, 0.2150)	(0.0057, 0.0779)	(0.0051, 0.0744)

Table 2 Empirical level of the test at nominal level 0.05 for Poisson INAR(1) model

λ	ϕ	Sample size				
		100	300	500	1000	5000
0.5	0.1	0.108	0.066	0.056	0.066	0.045
	0.3	0.081	0.078	0.046	0.064	0.052
	0.5	0.075	0.058	0.054	0.053	0.048
	0.7	0.072	0.063	0.069	0.057	0.053
	0.9	0.101	0.070	0.052	0.049	0.053
1	0.1	0.080	0.063	0.061	0.050	0.048
	0.3	0.078	0.048	0.055	0.050	0.051
	0.5	0.081	0.054	0.052	0.051	0.052
	0.7	0.077	0.061	0.059	0.051	0.049
	0.9	0.087	0.077	0.054	0.052	0.046
2	0.1	0.088	0.053	0.044	0.054	0.048
	0.3	0.070	0.066	0.052	0.056	0.051
	0.5	0.083	0.068	0.058	0.053	0.052
	0.7	0.061	0.065	0.056	0.058	0.047
	0.9	0.102	0.055	0.061	0.050	0.052

Vasconcellos (2015). Table 6 provides some descriptive statistics. The series and its sample autocorrelation function are shown in Fig. 1.

From mean, variance and correlation plots given in Fig. 1, it can be seen that the data can be modeled by a Poisson INAR(1) model and the constancy of the thinning parameter may be suspected. Therefore it is of interest to test the hypothesis of constancy of the thinning parameter. The value of the test statistic for testing $H_0 : \sigma_\phi^2 = 0$ against $H_1 : \sigma_\phi^2 > 0$ turned out to be 0.1035, with p -value 0.7477, indicating that one did not reject the null hypothesis at 5% level of significance. Consequently, it is better to model this data with a Poisson INAR(1) model rather than the RCINAR(1) model. Table 7 display the estimates. Two goodness-of-fit statistics are also shown: the root mean-squared error (RMSE) and mean absolute deviation (MAD). Those statistics defined as follows. For $t = 2, \dots, n$, consider the expected value of the observation at the previous time, $E[X_t|X_{t-1}] = \phi X_{t-1} + \lambda$. Let $\epsilon_t^* = X_t - \phi^* X_{t-1} - \lambda^*$, where ϕ^* and λ^* are any estimates of ϕ and λ . Define

$$RMSE = \sqrt{\frac{1}{n-1} \sum_{t=2}^n (\epsilon_t^*)^2}, \quad MAD = \frac{1}{n-1} \sum_{t=2}^n |\epsilon_t^*|.$$

From Table 7, it can be found that the CML and MELE are competitive and superior in terms of RMSE and MAD. We also plot the confidence region of the parameter as shown in Fig. 2. This is an indication that our proposed estimator can be indeed a good choice in practice.

We can also compare the estimators in terms of predictive power. Let ϕ_n^* and λ_n^* be any estimators of ϕ and λ , obtained from n observations. Then, at any given time n ,

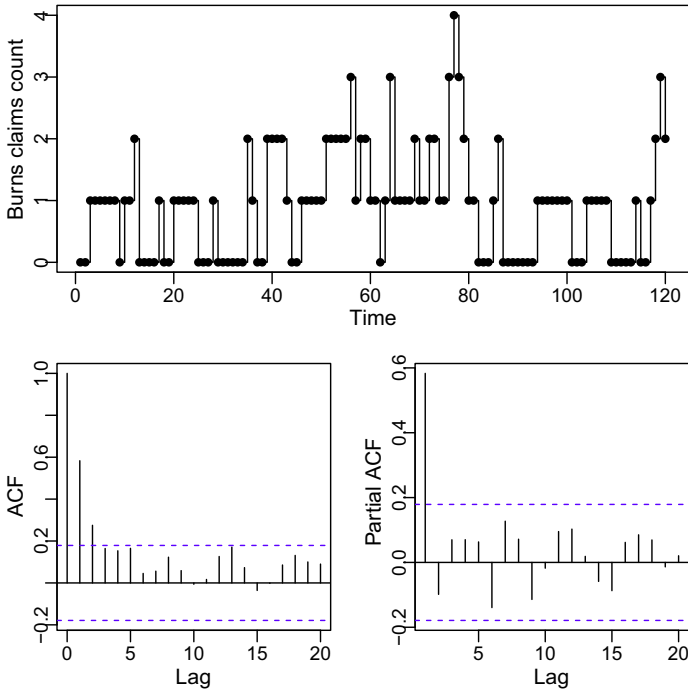


Fig. 1 The time plot, the sample autocorrelation and partial autocorrelation functions of the series of the burns claims

Fig. 2 Estimates and 95% confidence regions for (ϕ, λ) , based on CLS(Δ , dotted line), CML(\bullet , dashed line), EL($*$, solid line)

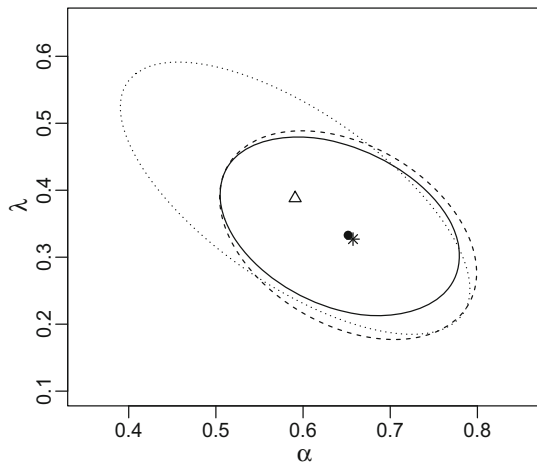


Table 3 Empirical power of the test at nominal level 0.05 for Poisson RCINAR(1) model with $\phi_t \sim \text{Beta}(a, b)$

λ	a	b	(ϕ, σ_ϕ^2)	Sample size				
				100	300	500	1000	5000
0.5	0.5	0.5	(0.500,0.125)	0.182	0.409	0.567	0.832	1.000
		1	(0.333,0.089)	0.120	0.146	0.181	0.358	0.989
		1.5	(0.250,0.063)	0.116	0.107	0.103	0.178	0.978
	1	0.5	(0.667,0.089)	0.178	0.441	0.680	0.920	1.000
		1	(0.500,0.083)	0.101	0.190	0.320	0.498	1.000
		1.5	(0.400,0.069)	0.092	0.120	0.178	0.294	0.985
	1.5	0.5	(0.750,0.063)	0.188	0.457	0.660	0.915	1.000
		1	(0.600,0.069)	0.118	0.228	0.313	0.541	1.000
		1.5	(0.500,0.063)	0.094	0.149	0.206	0.320	0.995
1	0.5	0.5	0.412	0.844	0.962	0.998	1.000	
		1	0.156	0.359	0.548	0.828	1.000	
		1.5	0.122	0.180	0.280	0.481	1.000	
	1	0.5	0.495	0.920	0.989	1.000	1.000	
		1	0.215	0.546	0.744	0.949	1.000	
		1.5	0.134	0.309	0.449	0.715	1.000	
	1.5	0.5	0.461	0.925	0.991	1.000	1.000	
		1	0.225	0.574	0.788	0.976	1.000	
		1.5	0.153	0.352	0.524	0.790	1.000	
2	0.5	0.5	0.841	0.999	1.000	1.000	1.000	
		1	0.402	0.814	0.967	0.999	1.000	
		1.5	0.197	0.432	0.698	0.926	1.000	
	1	0.5	0.913	1.000	1.000	1.000	1.000	
		1	0.559	0.943	0.999	1.000	1.000	
		1.5	0.313	0.707	0.898	0.994	1.000	
	1.5	0.5	0.907	1.000	1.000	1.000	1.000	
		1	0.601	0.973	0.999	1.000	1.000	
		1.5	0.364	0.808	0.945	0.999	1.000	

we can forecast the next observation as

$$X_{n+1}^* = \langle \phi_n^* X_n + \lambda_n^* \rangle,$$

where $\langle \cdot \rangle$ represents the nearest integer. The one-step ahead absolute forecast deviation is then given by

$$|X_{n+1}^* - X_{n+1}|,$$

where X_{n+1}^* is the one-step ahead forecast.

Table 4 Empirical level of the Awale et al. (2019) test at nominal level 0.05 for Poisson INAR(1) model

λ	ϕ	Sample size				
		100	300	500	1000	5000
0.5	0.1	0.058	0.062	0.055	0.065	0.055
	0.3	0.071	0.068	0.056	0.054	0.051
	0.5	0.080	0.068	0.064	0.052	0.049
	0.7	0.092	0.062	0.069	0.061	0.057
	0.9	0.121	0.090	0.062	0.059	0.055
1	0.1	0.085	0.073	0.065	0.060	0.046
	0.3	0.088	0.068	0.056	0.053	0.051
	0.5	0.082	0.064	0.062	0.055	0.054
	0.7	0.097	0.071	0.056	0.061	0.059
	0.9	0.107	0.078	0.074	0.058	0.048
2	0.1	0.098	0.063	0.074	0.057	0.049
	0.3	0.090	0.065	0.058	0.055	0.051
	0.5	0.097	0.078	0.068	0.043	0.053
	0.7	0.101	0.085	0.066	0.068	0.050
	0.9	0.132	0.095	0.068	0.065	0.047

Table 8 shows that all four estimators produce a correct forecast or a small forecast error in all the time interval here studied ($n = 111, \dots, 120$). This confirms that the PINAR(1) process is a good model for this data, at least in this interval. We observe that the CLS and MQL provide the worst forecasting performance. On the other hand, the MELE and CML provide the same forecasting and predict better than CLS and MQL. This is an indication that our proposed MELE estimator can be indeed a good choice for forecasting purposes.

5 Conclusions

In this article, we considered the empirical likelihood inference for the Poisson INAR(1) process by making effective use of auxiliary information and found a test statistic of testing the randomness of the parameter. A straightforward question is to ask this point is whether similar results as shown here could be obtained in the other models. This is our further work.

Funding This work is supported by National Natural Science Foundation of China (No. 11871028, 11731015, 11901053), Natural Science Foundation of Jilin Province (No. 20180101216JC).

6 Appendix

Proof of Theorem 1 The proof follows essentially the lines of the classical proof of Owen (2001) for the i.i.d. case, which may be applied when the following three conditions are satisfied. Let $S(\theta) = \frac{1}{n} \sum_{i=1}^n \mathbf{m}_i(\theta) \mathbf{m}_i(\theta)'$.

Table 5 Empirical power of the Awale et al. (2019) test at nominal level 0.05 for Poisson RCINAR(1) model with $\alpha_t \sim \text{Beta}(a, b)$

λ	a	b	(ϕ, σ_ϕ^2)	Sample size				
				100	300	500	1000	5000
0.5	0.5	0.5	(0.500,0.125)	0.382	0.909	0.999	1.000	1.000
		1	(0.333,0.089)	0.110	0.276	0.550	0.888	1.000
		1.5	(0.250,0.063)	0.056	0.157	0.153	0.578	0.982
	1	0.5	(0.667,0.089)	0.578	0.961	1.000	1.000	1.000
		1	(0.500,0.083)	0.191	0.610	0.790	0.968	1.000
		1.5	(0.400,0.069)	0.092	0.300	0.418	0.684	1.000
	1.5	0.5	(0.750,0.063)	0.688	0.987	1.000	1.000	1.000
		1	(0.600,0.069)	0.278	0.588	0.892	1.000	1.000
		1.5	(0.500,0.063)	0.090	0.249	0.554	0.782	0.995
1	0.5	0.5	0.712	0.984	1.000	1.000	1.000	
		1	0.256	0.559	0.848	0.978	1.000	
		1.5	0.102	0.250	0.421	0.711	1.000	
	1	0.5	0.815	0.980	1.000	1.000	1.000	
		1	0.365	0.876	0.744	1.000	1.000	
		1.5	0.214	0.409	0.449	0.965	1.000	
	1.5	0.5	0.860	1.000	1.000	1.000	1.000	
		1	0.525	0.934	1.000	1.000	1.000	
		1.5	0.223	0.672	0.834	0.990	1.000	
2	0.5	0.5	0.941	1.000	1.000	1.000	1.000	
		1	0.442	0.954	0.989	0.999	1.000	
		1.5	0.207	0.732	0.918	0.996	1.000	
	1	0.5	0.983	1.000	1.000	1.000	1.000	
		1	0.789	0.993	1.000	1.000	1.000	
		1.5	0.423	0.947	0.990	0.999	1.000	
	1.5	0.5	0.937	1.000	1.000	1.000	1.000	
		1	0.701	0.983	0.999	1.000	1.000	
		1.5	0.464	0.928	0.995	1.000	1.000	

Table 6 Descriptive statistic

Minimum	Median	Mean	Variance	Maximum
0	1	0.917	0.766	4

(C1) $S(\theta_0) \rightarrow \Sigma$ in probability and Σ positive definite.

(C2) $m_n^* = \max_{1 \leq t \leq n} \|m_t(\theta_0)\| = o_p(n^{1/2})$.

(C3) $\frac{1}{n} \sum_{t=1}^n \|m_t(\theta_0)\|^3 = o_p(n^{1/2})$, where θ_0 is the true value of the parameter.

Table 7 Estimates of the parameters (standard errors in parentheses), RMSE and MAD for burns claims

Method	Parameter		RMSE	MAD
	$\hat{\phi}$	$\hat{\lambda}$		
CLS	0.5909 (0.0819)	0.3881 (0.0830)	0.7065	0.5248
MQL	0.5981 (0.0810)	0.3816 (0.0802)	0.7065	0.5238
CML	0.6517 (0.0602)	0.3329 (0.0636)	0.7055	0.5164
MELE	0.6573 (0.0616)	0.3270 (0.0595)	0.7058	0.5158

Table 8 $|X_{n+1}^* - X_{n+1}|$ for counts of burns claim

n	CLS	MQL	MELE	CML
111	0	0	0	0
112	0	0	0	0
113	1	1	0	0
114	0	0	0	0
115	0	0	0	0
116	1	1	1	1
117	0	0	0	0
118	0	0	0	0
119	1	1	1	1
120	1	1	1	1

For the first condition, let $\mathcal{F}_n = \sigma(X_0, X_1, \dots, X_n)$, $A_n = \sum_{t=1}^n m_{1t}(\theta_0) = \sum_{t=1}^n (X_t - \phi_0 X_{t-1} - \lambda_0)$, $A_0 = 0$, then,

$$E(A_n | \mathcal{F}_{n-1}) = A_{n-1} + E(X_n - \phi_0 X_{n-1} - \lambda_0) = A_{n-1},$$

so, $\{A_n, \mathcal{F}_n, n \geq 0\}$ is a martingale. For $E|X_t|^2 < \infty$, $\{(X_t - \phi_0 X_{t-1} - \lambda_0)^2, n \geq 1\}$ is square integrable. By the strictly stationary and ergodic theorem,

$$\frac{1}{n} \sum_{t=1}^n (X_t - \phi_0 X_{t-1} - \lambda_0)^2 \xrightarrow{a.s.} E(X_1 - \phi_0 X_0 - \lambda_0)^2.$$

Similarly, we can proof that $B_n = \sum_{t=1}^n m_{2t}(\theta_0)$, $C_n = \sum_{t=1}^n m_{3t}(\theta_0)$ are martingale. Then the first condition holds. To check the second condition, we note that $E\{m_t(\theta_0)/m_t(\theta_0)\} < \infty$, then, $\sum_{n=1}^{\infty} P(m_t(\theta_0)' m_t(\theta_0) > n) < \infty$. By the Borel-Cantelli lemma, $\|m_t(\theta_0)\| > \sqrt{n}$ happens with probability 1 only for finitely many n , since $\{X_t\}$ is a strictly stationary process. This implies that there are only finitely many n for which $m_n^* > \sqrt{n}$. Similarly, for any $\varepsilon > 0$, there are only finitely many n for which $m_n^* > \varepsilon\sqrt{n}$, hence $\limsup_{n \rightarrow \infty} m_n^*/\sqrt{n} \leq \varepsilon$ holds with probability 1, so $m_n^* = o_p(\sqrt{n})$. From the (C1) and (C2), we have $\frac{1}{n} \sum_{t=1}^n \|m_t(\theta_0)\|^3 \leq m_n^* \frac{1}{n} \sum_{t=1}^n m_t(\theta_0)m_t(\theta_0)' = o_p(\sqrt{n})O_p(1) = o_p(n^{1/2})$. See Zhang et al. (2011a) and Zhao and Yu (2016) also.

Proof of Theorem 2 The proof is similar to that of Theorem 6.2 of Zhang et al. (2011a) and that of Theorem 1 of Qin and Lawless (1994). Let

$$Q_{1n}(\theta, \beta) = \frac{1}{n} \sum_{t=1}^n \frac{m_t(\theta)}{1 + \beta' m_t(\theta)}, \quad Q_{2n}(\theta, \beta) = \frac{1}{n} \sum_{t=1}^n \frac{\beta(\theta)}{1 + \beta' m_t(\theta)} \left(\frac{\partial m_t(\theta)}{\partial \theta} \right)'.$$

Expanding $Q_{1n}(\check{\theta}, \check{\beta}), Q_{2n}(\check{\theta}, \check{\beta})$ at $(\theta_0, 0)$, we have

$$0 = Q_{1n}(\check{\theta}, \check{\beta}) = Q_{1n}(\theta_0, 0) + \frac{\partial Q_{1n}(\theta_0, 0)}{\partial \theta} (\check{\theta} - \theta_0) + \frac{\partial Q_{1n}(\theta_0, 0)}{\partial \beta'} (\check{\beta} - 0) + o_p(\delta_n),$$

$$0 = Q_{2n}(\check{\theta}, \check{\beta}) = Q_{2n}(\theta_0, 0) + \frac{\partial Q_{2n}(\theta_0, 0)}{\partial \theta} (\check{\theta} - \theta_0) + \frac{\partial Q_{2n}(\theta_0, 0)}{\partial \beta'} (\check{\beta} - 0) + o_p(\delta_n),$$

where $\delta_n = \|\check{\theta} - \theta_0\| + \|\check{\beta}\| = O_p(n^{-1/2})$. So, we have

$$S_n \begin{pmatrix} \check{\beta} \\ \check{\theta} - \theta_0 \end{pmatrix} = \begin{pmatrix} -Q_{1n}(\theta_0, 0) + o_p(\delta_n) \\ o_p(\delta_n) \end{pmatrix}.$$

where

$$S_n = \begin{pmatrix} \frac{\partial Q_{1n}}{\partial \beta'} & \frac{\partial Q_{1n}}{\partial \theta} \\ \frac{\partial Q_{2n}}{\partial \beta'} & 0 \end{pmatrix}_{(\theta_0, 0)} \rightarrow \begin{pmatrix} -E(m_t m_t') & E \left(\frac{\partial m_t}{\partial \theta} \right) \\ E \left(\frac{\partial m_t}{\partial \theta} \right)' & 0 \end{pmatrix} = \begin{pmatrix} -\Sigma & \Omega \\ \Omega' & 0 \end{pmatrix}.$$

since $\{X_t\}$ is strictly stationary and ergodic and note that $\sqrt{n}Q_{1n}(\theta_0, 0) = \frac{1}{\sqrt{n}} \sum_{t=1}^n m_t(\theta_0) \xrightarrow{d} N(0, \Sigma)$. Then, we have

$$\sqrt{n}(\check{\theta} - \theta_0) = (\Omega' \Sigma^{-1} \Omega)^{-1} \Omega' \Sigma^{-1} \sqrt{n}Q_{1n}(\theta_0, 0) + o_p(1).$$

The proof of the theorem is completed.

Proof of Theorem 3: The proof is similar to that of Theorem 2 and Corollary 4 of Qin and Lawless (1994) and will be only sketched here.

Note that

$$\ell(\check{\theta}, \check{\beta}) = \sum_{t=1}^n \log[1 + \check{\beta}' m_t(\check{\theta})] = \frac{n}{2} \left[\frac{1}{n} \sum_{t=1}^n m_t(\theta_0)' \right] A \left[\frac{1}{n} \sum_{t=1}^n m_t(\theta_0) \right] + o_p(1),$$

where $A = \Sigma^{-1} [I - \Omega [\Omega' \Sigma^{-1} \Omega]^{-1} \Omega' \Sigma^{-1}]$. Thus,

$$2\ell(\check{\theta}) = [\Sigma^{-1/2} \sqrt{n}Q_{1n}(\theta_0, 0)]' [I - \Sigma^{-1/2} \Omega [\Omega' \Sigma^{-1} \Omega]^{-1} \Omega' \Sigma^{-1/2}] [\Sigma^{-1/2} \sqrt{n}Q_{1n}(\theta_0, 0)] + o_p(1),$$

Note that $(\Sigma)^{-1/2} \sqrt{n} Q_{1n}(\theta_0, 0)$ converges to a standard multivariate normal distribution and the matrix $I - \Sigma^{-1/2} \Omega [\Omega' \Sigma^{-1} \Omega]^{-1} \Omega' \Sigma^{-1/2}$ is symmetric and idempotent, with trace equal to 1. Hence the test statistic $2\ell(\check{\theta})$ converges to $\chi^2(1)$.

References

- Al-Osh MA, Aly EEAA (1992) First order autoregressive time series with negative binomial and geometric marginals. *Commun Stat Theory Methods* 21:2483–2492
- Al-Osh MA, ALZAID AA (1987) First-order integer-valued autoregressive (INAR(1)) process. *J Time Series Anal* 8:261–275
- Alzaid A, Al-Osh M (1988) First-Order Integer-Valued Autoregressive (INAR (1)) Process: Distributional and Regression Properties. *Stat Neerlandica* 42:53–61
- Awale M, Balakrishna N, Ramanathan TV (2019) Testing the constancy of the thinning parameter in a random coefficient integer autoregressive model. *Stat Papers* 60:1515–1539
- Bourguignon M, Vasconcellos KLP (2015) Improved estimation for poisson INAR(1) models. *J Statal Comput Simul* 85:2425–2441
- Freeland RK, McCabe BPM (2004) Analysis of low count time series data by Poisson autoregression. *J Time Series Anal* 25:701–722
- Kang J, Lee S (2009) Parameter change test for random coefficient integer-valued autoregressive processes with application to polio data analysis. *J Time Series Anal* 30:239–258
- Keith Freeland R, McCabe B (2005) Asymptotic properties of CLS estimators in the Poisson AR(1) model. *Stat Probab Lett* 73:147–153
- Owen A (1991) Empirical likelihood for linear models. *Annals Stat* 19:1725–1747
- Owen AB (1988) Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* 75:237–249
- Owen AB (2001) *Empirical likelihood*. Chapman and Hall, London
- Qin J, Lawless J (1994) Empirical likelihood and general estimating equations. *Annals Stat* 22:300–325
- Qin J, Lawless J (1995) Estimating equations, empirical likelihood and constraints on parameters. *Canadian J Stat / La Revue Canadienne de Statistique* 23:145–159
- Steutel F, Van Harn K (1979) Discrete analogues of self-decomposability and stability. *Annals Probab* 74:893–899
- Weiss CH (2017) *An introduction to discrete-valued time series*. John Wiley & Sons, Hoboken, NJ
- Zhang H, Wang D, Zhu F (2011a) The empirical likelihood for first-order random coefficient integer-valued autoregressive processes. *Commun Stat Theory Methods* 40:492–509
- Zhang H, Wang D, Zhu F (2011) Empirical likelihood inference for random coefficient INAR(p) process. *J Time Series Anal* 32:195–203
- Zhao Z, Hu Y (2015) Statistical inference for first-order random coefficient integer-valued autoregressive processes. *J Inequal Appl* 2015:359
- Zhao Z, Yu W (2016) Empirical Likelihood Inference for First-Order Random Coefficient Integer-Valued Autoregressive Processes. *Math Probl Eng* 2016:1–8
- Zheng H, Basawa IV, Datta S (2007) First-order random coefficient integer-valued autoregressive processes. *J Stat Plann Inference* 137:212–229

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.