RESEARCH

Journal of Inequalities and Applications a SpringerOpen Journal

Open Access



Empirical likelihood-based inference in Poisson autoregressive model with conditional moment restrictions

Cuixin Peng^{1,2}, Dehui Wang^{1*} and Zhiwen Zhao³

*Correspondence: wangdh@jlu.edu.cn ¹College of Mathematics, Jilin University, Changchun, 130012, China Full list of author information is available at the end of the article

Abstract

In this paper, we employ the empirical likelihood method to estimate the unknown parameters in Poisson autoregressive model in the presence of auxiliary information. It is shown that our approach proposed, compared to the maximum likelihood estimator, the least squares estimator, and the weighted least squares estimator, yields more efficient estimators. Some simulation studies are also conducted to investigate the finite sample performance of the proposed method.

MSC: Primary 62M10; secondary 62M05

Keywords: conditional moment restriction; empirical likelihood; least square estimator; asymptotic distribution; Poisson autoregressive model

1 Introduction

Integer-valued time series occur in many different situations. They arise, for example, as the number of births at a hospital in successive months, the number of bases of DNA sequences, the number of road accidents and the number of diseases in a certain area in successive months. Therefore, in recent years, there has been growing interest in study-ing integer-valued time series [1–4]. In order to model the number of cases of campy-lobacterosis infections from January 1990 to the end of October 2000 in the north of the Province of Québec, Ferland *et al.* [5] proposed the following Poisson autoregressive model:

$$\begin{cases} X_t | \mathscr{F}_{t-1} : \mathcal{P}(\lambda_t); & \forall t \in \mathbb{Z}, \\ \lambda_t = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}, \end{cases}$$
(1.1)

where \mathscr{F}_{t-1} is the σ field generated by $(X_{t-1}, X_{t-2}, ...), \alpha_0 > 0, \alpha_i \ge 0$ (i = 1, 2, ..., p), and $\alpha = (\alpha_0, \alpha_1, ..., \alpha_p)$ is an unknown parameter vector.

For model (1.1), Zhu and Wang [6] gave the condition for ergodicity and a necessary and sufficient condition for the existence of moments. They also established the asymptotics for the maximum likelihood estimator and the least squares estimators. The problem of interest here is to estimate the unknown parameter in model (1.1) by using the empirical likelihood method when auxiliary information is available. In practice, some auxiliary information can often be obtained, such as that the unknown distribution is symmetric or



© 2015 Peng et al. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

the variance of population is a function of the mean. By making full use of the auxiliary information, we can increase the precision of statistical inference. Here, we assume that we have some auxiliary information that can be represented as the following conditional moment restrictions:

$$E(g(X_t, \dots, X_{t-p}; \theta_0) | X(t-1)) = 0$$
(1.2)

for each t = 0, 1, 2, ..., where the unknown parameter vector $\theta_0 \in \mathbb{R}^d$, $X(t - 1) = (X_{t-1}, ..., X_{t-p})$ and $g(x; \theta) \in \mathbb{R}^r$ is some function with $r \ge d$. In order to simplify the notation, we further denote $g(X_t, ..., X_{t-p}; \theta)$ by $g_t(\theta)$. We note here that θ can be different from α and the notion $g_t(\theta)$ contains a broad class of information that can be formulated from the knowledge on the probability distribution of X_t , *e.g.* the moment and their generalizations [7]. By using the conditional moment restrictions, as we expect, we can increase the efficiency of the resulting estimator [8–10].

The EL as an alternative to the bootstrap for constructing confidence regions was introduced by Owen [11, 12]. The method defines an EL ratio function to construct confidence regions. Important features of the empirical likelihood method are its automatic determination of the shape and orientation of the confidence region by the data. These attractive properties have motivated various authors to extend the empirical likelihood methodology to other situations. To use the auxiliary information, some statisticians have also developed some statistical inference methods under the framework of empirical likelihood method [13–15]. In this paper, we further generalize these methods to the statistical inference of time series models. Specifically, based on the empirical likelihood method, we consider the parameter estimation problem for Poisson autoregressive model with conditional moment restrictions. Our approach yields more efficient estimates compared to the maximum likelihood estimator, the least squares estimator and the weighted least squares estimator, which do not utilize the conditional moment restrictions. Based on the mean square errors, a comparison is also made by simulation. Our simulation indicates that the use of auxiliary information provides improved inferences.

The rest of this paper is organized as follows. In Section 2, we introduce the methodology and the main results. Simulation results are reported in Section 3. Section 4 provides the proofs of the main results.

The symbols $\stackrel{d}{\longrightarrow}$ and $\stackrel{p}{\longrightarrow}$ denote convergence in distribution and convergence in probability, respectively. Convergence 'almost surely' is written as 'a.s.'. Furthermore, ' $M_{k\times p}^{\tau}$ ' denotes the transpose matrix of the $k \times p$ matrix $M_{k\times p}$, $A \otimes B$ denotes the Kronecker product of matrices A and B, and $\|\cdot\|$ denotes Euclidean norm of the matrix or vector.

2 Methodology and main results

In this section, we will first discuss how to apply the empirical likelihood method [16, 17] to estimate the unknown parameter α when auxiliary information is available.

Before we state our main results, the following assumptions will be made:

(A₁) The parametric space Υ is compact with $\Upsilon = \{\alpha : \delta \le \alpha_0 \le M, 0 \le \alpha_1 + \dots + \alpha_p \le M^* < 1, \alpha_i \ge 0, i = 1, 2, \dots, p\}$, where δ and M are finite positive constants, and the true parameter value α^0 is an interior point in Υ .

Remark 1 It is shown by Corollary 1 in Ferland *et al.* [5] and Theorem 1 and Theorem 2 in Zhu and Wang [6] that, under (A₁), $\{X_t, t \ge 1\}$ is stationary and ergodic, and $E(X_t^m) < \infty$ for any fixed positive integer *m*.

(A₂) There exists θ_0 such that $E(g_t(\theta_0)) = 0$, the matrix $\Sigma(\theta) = E(g_t(\theta)g_t^{\mathsf{T}}(\theta))$ is positive definite at θ_0 , $\partial g(x;\theta)/\partial \theta$ is continuous in a neighborhood of the true value θ_0 , $\|\partial g(x;\theta)/\partial \theta\|$ and $\|g(x;\theta)\|^3$ are bounded by some integrable function $\tilde{W}(x)$ in this neighborhood, and the rank of $E(\partial g_t(\theta)/\partial \theta)$ is *d*.

First, the conditional moment restrictions in (1.2) imply that $E(g_t(\theta_0)) = 0$. Further, by using the empirical likelihood method, we can obtain data adaptive weights ω_t through

$$L(\theta) = \sup\left\{\prod_{t=1}^{n} \omega_t : \omega_t \ge 0, \sum_{i=1}^{n} \omega_t = 1, \sum_{t=1}^{n} \omega_t g_t(\theta_0) = 0\right\},\$$

where θ_0 is an unknown parameter. By using the auxiliary information combining with the least squares method, we propose to estimate α by

$$\hat{\alpha} = \arg\min_{\alpha} \sum_{t=1}^{n} \omega_t (X_t - Z_t^{\tau} \alpha)^2, \qquad (2.1)$$

where $Z_t^{\tau} = (1, X_{t-1}, \dots, X_{t-p})$. By introducing a Lagrange multiplier $\lambda \in \mathbb{R}^r$, standard derivations in the empirical likelihood lead to

$$\omega_t(\theta_0) = \frac{1}{n} \frac{1}{1 + \lambda_{\theta_0}^{\tau} g_t(\theta_0)},$$
(2.2)

where λ_{θ_0} satisfies

$$\frac{1}{n} \sum_{t=1}^{n} \frac{g_t(\theta_0)}{1 + \lambda_{\theta_0}^{\tau} g_t(\theta_0)} = 0.$$
(2.3)

Utilizing the weights given by (2.2), we obtain the estimate of α :

$$\hat{\alpha} = \left(\sum_{t=1}^{n} \omega_t(\theta_0) Z_t Z_t^{\tau}\right)^{-1} \sum_{t=1}^{n} \omega_t(\theta_0) X_t Z_t.$$
(2.4)

In the following, we will give the asymptotic properties of $\hat{\alpha}$.

Theorem 2.1 Assume that (A_1) and (A_2) hold. If α^0 is the true value of α , then

$$\sqrt{n}(\hat{\alpha} - \alpha^0) \stackrel{d}{\longrightarrow} N(0, W^{-1}(\Lambda - \Lambda_{12}\Sigma^{-1}(\theta_0)\Lambda_{12}^{\tau})W^{-1}),$$

where $W = E(Z_t Z_t^{\tau})$, $\Lambda = E(Z_t Z_t^{\tau} (X_t - Z_t^{\tau} \alpha^0)^2)$, and $\Lambda_{12} = E(Z_t g_t^{\tau} (\theta_0) (X_t - Z_t^{\tau} \alpha^0))$.

Zhu and Wang [6] prove that the asymptotic variance of the ordinary least squares estimator $\bar{\alpha} = (\sum_{t=1}^{n} Z_t Z_t^{\tau})^{-1} \sum_{t=1}^{n} X_t Z_t$ is $W^{-1} \Lambda W^{-1}$. Note that Λ and $\Sigma^{-1}(\theta_0)$ are both positive definite matrices. Therefore, by Theorem 2.1, we find that a variance reduction

quantified by $W^{-1}\Lambda_{12}\Sigma^{-1}(\theta_0)\Lambda_{12}^{\tau}W^{-1}$ is induced by incorporating the auxiliary information $E(g_t(\theta_0)|X(t-1)) = 0$.

To apply the proposed estimator (2.4), we need to further estimate the unknown parameter θ . We consider $\hat{\theta} = \arg \max_{\theta} L(\theta)$. Following the results in Qin and Lawless [16], the corresponding weights $\{\omega_t\}_{t=1}^n$ satisfy

$$\omega_t(\hat{\theta}) = \frac{1}{n} \frac{1}{1 + \lambda_{\hat{\alpha}}^{\tau} g_t(\hat{\theta})},\tag{2.5}$$

where $\lambda_{\hat{\theta}}$ is the solution to

$$\frac{1}{n}\sum_{t=1}^{n}\frac{g_t(\hat{\theta})}{1+\lambda_{\hat{\theta}}^{\tau}g_t(\hat{\theta})}=0$$

and $(\lambda_{\hat{\theta}}, \hat{\theta})$ solves

$$\frac{1}{n}\sum_{t=1}^{n}\frac{\frac{\partial g_{t}(\theta)}{\partial \theta^{\tau}}\lambda_{\hat{\theta}}}{1+\lambda_{\hat{\theta}}^{\tau}g_{t}(\hat{\theta})}=0$$

Let

$$\hat{\alpha}^{1} = \arg\min_{\alpha} \sum_{t=1}^{n} \omega_{t}(\hat{\theta}) \left(X_{t} - Z_{t}^{\tau} \alpha \right)^{2},$$
(2.6)

where $\{\omega_t(\hat{\theta})\}_{t=1}^n$ is identified by (2.5). When r = d, we know that $\omega_t(\hat{\theta}) = \frac{1}{n}$ and hence $\hat{\alpha}^1$ is the ordinary least squares estimator. When r > d, $\omega_t(\hat{\theta})$ is no longer equal to $\frac{1}{n}$ and we shall show that this scheme provides an efficiency gain over the conventional least squares estimator.

In order to study the estimator (2.6), we define $\Gamma(\theta_0) = E(\frac{\partial g_I(\theta_0)}{\partial \theta})$, $\Omega(\theta_0) = (\Gamma(\theta_0)\Sigma^{-1}(\theta_0) \times \Gamma^{\tau}(\theta_0))^{-1}$ and $B = \Sigma^{-1}(\theta_0)(I - \Gamma(\theta_0)\Omega(\theta_0)\Gamma^{\tau}(\theta_0)\Sigma^{-1}(\theta_0))$, where *I* is the identity matrix. The limiting distribution of $\hat{\alpha}^1$ is given in the following theorem.

Theorem 2.2 Assume that (A_1) and (A_2) hold. If α^0 is the true value of α , then

$$\sqrt{n}(\hat{\alpha}^{1} - \alpha^{0}) \xrightarrow{d} N(0, W^{-1}(\Lambda - \Lambda_{12}B\Lambda_{12}^{\tau})W^{-1}).$$
(2.7)

The matrix *B* is non-negative definite. Hence the asymptotic variance of $\hat{\alpha}^1$ is no greater than that of the least squares estimator. When *B* is positive definite, variance reduction is attained. This implies that having much more auxiliary information can improve the least squares estimator.

3 Simulation study

In this section we conduct some simulation studies which show that our proposed methods perform very well. Consider the following one order Poisson autoregressive model:

$$\begin{cases} X_t | \mathscr{F}_{t-1} : \mathcal{P}(\lambda_t); & \forall t \in \mathbb{Z}, \\ \lambda_t = \alpha_0 + \alpha_1 X_{t-1}. \end{cases}$$
(3.1)

In order to compare the performance of the estimator (denoted by ALS) given by (2.6) with those of the ordinary least squares estimator (LS), the weighted least squares estimator (WLS), and the maximum likelihood estimation (MLE), we compute the mean square errors based on the four methods: the ALS, the LS, the WLS, and the MLE. We use the vector $g_t(\alpha_0, \alpha_1) = (1, X_{t-1}, X_{t-1}^2)^{\tau} (X_t - \alpha_0 - \alpha_1 X_{t-1})$ as the conditional moment restrictions in (1.2). Specifically, for a particular pair of $(\alpha_0, \alpha_1)^{\tau}$, we generate realizations from (3.1) with n = 100,300 and 1,000. Further, based on 1,000 repetitions, we compute the mean square errors of the above four kinds of estimators. The simulation results for $\alpha_0 = 1$ are summarized in Table 1. Table 2 presents the simulation results for $\alpha_0 = 2$.

From Tables 1 and 2, we see that the mean square errors obtained by the estimator (2.6) are less than those of the maximum likelihood estimator, the least squares estimator and the weighted least squares estimator. This indicates that using the conditional moment restrictions, the estimates are more accurate, regardless of the samples size and the different unknown parameter.

4 Proofs of the main results

In order to prove Theorem 2.1, we first present several lemmas.

Lemma 4.1 Assume that (A_1) and (A_2) hold. Then

$$\max_{1 \le t \le n} \left\| g_t(\theta_0) \right\| = o_p(n^{\frac{1}{2}}).$$
(4.1)

Proof Let $\Sigma_n(\theta_0) = \frac{1}{n} \sum_{t=1}^n (g_t(\theta_0) g_t^{\tau}(\theta_0))$ and $g_t(\theta_0) = (g_{t1}(\theta_0), \dots, g_{tr}(\theta_0))^{\tau}$. Note that in order to prove (4.1), we need only to prove that

$$\frac{1}{n} \max_{1 \le t \le n} g_{tk}(\theta_0) \xrightarrow{p} 0, \quad k = 1, \dots, r.$$
(4.2)

We denote the (h, l)th element of $\Sigma(\theta_0)$ as σ_{hl} , h, l = 1, 2, ..., r. By the ergodic theorem, we have

$$\Sigma_n(\theta_0) \xrightarrow{P} \Sigma(\theta_0). \tag{4.3}$$

For all $1 \le k \le r$ and $1 \le m \le n$, define the sets

$$B_{n,m}^k = \bigcap_{j=1}^m B_{n,m}^{k,j},$$

where $B_{n,m}^{k,j} = \{\omega : |\frac{1}{n} \sum_{t=1}^{[n(j/m)]} g_{tk}^2(\theta_0) - \frac{j}{m} \sigma_{kk}| \le \frac{1}{m}\}, j = 1, ..., m.$ First we show that

$$\lim_{n \to \infty} P(B_{n,m}^{k,j}) = 1, \quad j = 1, \dots, m.$$
(4.4)

After some algebra, we obtain

$$B_{n,m}^{k,j} = \left\{ \omega : \left| \frac{[n\frac{j}{m}]}{n} \frac{m}{j} \frac{1}{[n\frac{j}{m}]} \sum_{t=1}^{[n(j/m)]} g_{tk}^2(\theta_0) - \sigma_{kk} \right| \le \frac{1}{m} \frac{m}{j} \right\}.$$

	(α_0, α_1)	ALS	LS	WLS	MLE
n = 100	(1, 0.1)	(0.0226, 0.0085)	(0.0223, 0.0079)	(0.0231, 0.0081)	(0.0452, 0.0169)
	(1, 0.2)	(0.0242, 0.0104)	(0.0240, 0.0108)	(0.0279, 0.0127)	(0.0485, 0.0208)
	(1, 0.3)	(0.0250, 0.0103)	(0.0261, 0.0107)	(0.0310, 0.0126)	(0.0501, 0.0206)
	(1, 0.4)	(0.0145, 0.0067)	(0.0175, 0.0076)	(0.0275, 0.0099)	(0.0808, 0.0626)
	(1, 0.5)	(0.0283, 0.0098)	(0.0333, 0.0102)	(0.0647, 0.0145)	(0.1811, 0.1731)
	(1, 0.6)	(0.0673, 0.0120)	(0.0774, 0.0145)	(0.1235, 0.0207)	(0.1580, 0.0577)
	(1, 0.7)	(0.1085, 0.0090)	(0.1550, 0.0123)	(0.2932, 0.0196)	(0.2232, 0.0265)
	(1, 0.8)	(0.1500, 0.0061)	(0.2162, 0.0079)	(0.4171, 0.0123)	(0.3046, 0.0152)
	(1, 0.9)	(0.3878, 0.0046)	(0.3945, 0.0055)	(0.9465, 0.0102)	(0.6821, 0.0089)
	(1, 0.95)	(2.5205, 0.0054)	(3.1072, 0.0064)	(5.3794, 0.0109)	(4.7175, 0.0104)
n = 300	(1, 0.1)	(0.0071, 0.0032)	(0.0072, 0.0033)	(0.0077, 0.0036)	(0.0142, 0.0065)
	(1, 0.2)	(0.0092, 0.0041)	(0.0094, 0.0044)	(0.0105, 0.0050)	(0.0183, 0.0082)
	(1, 0.3)	(0.0087, 0.0034)	(0.0094, 0.0037)	(0.0128, 0.0049)	(0.0174, 0.0067)
	(1, 0.4)	(0.0102, 0.0041)	(0.0119, 0.0045)	(0.0191, 0.0060)	(0.0204, 0.0082)
	(1, 0.5)	(0.0113, 0.0032)	(0.0128, 0.0036)	(0.0238, 0.0053)	(0.0226, 0.0065)
	(1, 0.6)	(0.0169, 0.0027)	(0.0195, 0.0034)	(0.0452, 0.0062)	(0.0337, 0.0054)
	(1, 0.7)	(0.0188, 0.0016)	(0.0261, 0.0023)	(0.0723, 0.0046)	(0.0376, 0.0032)
	(1, 0.8)	(0.0307, 0.0014)	(0.0499, 0.0021)	(0.0146, 0.0046)	(0.0615, 0.0028)
	(1, 0.9)	(0.0536, 0.0009)	(0.0881, 0.0012)	(0.2846, 0.0024)	(0.1071, 0.0018)
	(1, 0.95)	(0.2441, 0.0006)	(0.2555, 0.0007)	(0.6927, 0.0014)	(0.4882, 0.0013)
n = 1,000	(1, 0.1)	(0.0018, 0.0009)	(0.0017, 0.0009)	(0.0018, 0.0010)	(0.0035, 0.0018)
	(1, 0.2)	(0.0025, 0.0011)	(0.0025, 0.0011)	(0.0027, 0.0013)	(0.0050, 0.0022)
	(1, 0.3)	(0.0029, 0.0012)	(0.0033, 0.0013)	(0.0046, 0.0018)	(0.0058, 0.0023)
	(1, 0.4)	(0.0027, 0.0010)	(0.0029, 0.0012)	(0.0055, 0.0019)	(0.0055, 0.0021)
	(1, 0.5)	(0.0030, 0.0007)	(0.0036, 0.0009)	(0.0083, 0.0017)	(0.0061, 0.0013)
	(1, 0.6)	(0.0034, 0.0007)	(0.0045, 0.0008)	(0.0109, 0.0014)	(0.0068, 0.0014)
	(1, 0.7)	(0.0077, 0.0008)	(0.0090, 0.0010)	(0.0270, 0.0020)	(0.0155, 0.0016)
	(1, 0.8)	(0.0095, 0.0005)	(0.0140, 0.0008)	(0.0428, 0.0015)	(0.0191, 0.0010)
	(1, 0.9)	(0.0211, 0.0003)	(0.0281, 0.0004)	(0.0854, 0.0007)	(0.0421, 0.0005)
	(1, 0.95)	(0.0244, 0.0001)	(0.0518, 0.0001)	(0.2280, 0.0004)	(0.0488, 0.0002)

Table 1 Simulation results when $\alpha_0 = 1$

Observe that

$$\frac{[n\frac{j}{m}]}{n}\frac{m}{j} \to 1, \quad n \to \infty.$$
(4.5)

Furthermore, by (4.3), we have

$$\frac{1}{[n_m^j]} \sum_{t=1}^{[n(j/m)]} g_{tk}^2(\theta_0) \xrightarrow{p} \sigma_{kk}.$$
(4.6)

Using (4.5) and (4.6), we complete the proof of (4.4). Next we prove that

$$\lim_{n \to \infty} P(B_{n,m}^k) = 1, \quad m = 1, 2, \dots$$
(4.7)

After some calculation, we obtain

$$P(B_{n,m}^{k}) = P\left(\bigcap_{j=1}^{m} B_{n,m}^{k,j}\right)$$
$$= P\left(\bigcap_{j=1}^{m-1} B_{n,m}^{k,j}\right) + P(B_{n,m}^{k,m}) - P\left(\left(\bigcap_{j=1}^{m-1} B_{n,m}^{k,j}\right) \cup B_{n,m}^{k,m}\right)$$

	(α_0, α_1)	ALS	LS	WLS	MLE
n = 100	(2, 0.1)	(0.0672, 0.0113)	(0.0658, 0.0108)	(0.0665, 0.0106)	(0.1762, 0.0479)
	(2, 0.2)	(0.0751, 0.0086)	(0.0752, 0.0087)	(0.0797, 0.0095)	(0.5044, 0.1226)
	(2, 0.3)	(0.0951, 0.0103)	(0.0986, 0.0106)	(0.1146, 0.0121)	(0.1908, 0.0298)
	(2, 0.4)	(0.1464, 0.0124)	(0.1591, 0.0133)	(0.1914, 0.0154)	(0.2929, 0.0321)
	(2, 0.5)	(0.1706, 0.0097)	(0.1826, 0.0105)	(0.2329, 0.0132)	(0.3307, 0.0209)
	(2, 0.6)	(0.2284, 0.0096)	(0.2546, 0.0110)	(0.3457, 0.0140)	(0.4378, 0.0191)
	(2, 0.7)	(0.2862, 0.0066)	(0.3291, 0.0077)	(0.4491, 0.0101)	(0.5638, 0.0132)
	(2, 0.8)	(0.5566, 0.0065)	(0.7199, 0.0084)	(1.0957, 0.0119)	(1.1036, 0.0358)
	(2, 0.9)	(2.0644, 0.0042)	(2.3995, 0.0051)	(3.4865, 0.0075)	(4.1109, 0.0082)
	(2, 0.95)	(10.19088, 0.0049)	(11.6877, 0.0057)	(15.3550, 0.0076)	(20.3388, 0.0098)
n = 300	(2, 0.1)	(0.0200, 0.0032)	(0.0202, 0.0032)	(0.0211, 0.0033)	(1.3292, 0.3556)
	(2, 0.2)	(0.0288, 0.0041)	(0.0287, 0.0042)	(0.0300, 0.0044)	(0.1811, 0.0696)
	(2, 0.3)	(0.0383, 0.0037)	(0.0403, 0.0037)	(0.0456, 0.0041)	(0.0755, 0.0134)
	(2, 0.4)	(0.0379, 0.0029)	(0.0419, 0.0034)	(0.0532, 0.0043)	(0.0711, 0.0033)
	(2, 0.5)	(0.0548, 0.0035)	(0.0624, 0.0039)	(0.0796, 0.0048)	(0.1052, 0.0075)
	(2, 0.6)	(0.0612, 0.0026)	(0.0779, 0.0033)	(0.1276, 0.0049)	(0.1194, 0.0053)
	(2, 0.7)	(0.0824, 0.0023)	(0.0961, 0.0025)	(0.1537, 0.0034)	(0.1613, 0.0045)
	(2, 0.8)	(0.1095, 0.0014)	(0.1194, 0.0015)	(0.1929, 0.0022)	(0.2158, 0.0027)
	(2, 0.9)	(0.3058, 0.0010)	(0.4142, 0.0013)	(0.8087, 0.0022)	(0.6160, 0.0020)
	(2, 0.95)	(1.0248, 0.0007)	(1.1210, 0.0008)	(1.7182, 0.0012)	(2.0500, 0.0015)
<i>n</i> = 1,000	(2, 0.1)	(0.0063, 0.0009)	(0.0063, 0.0009)	(0.0064, 0.0009)	(1.0834, 0.2918)
	(2, 0.2)	(0.0081, 0.0010)	(0.0084, 0.0010)	(0.0091, 0.0011)	(0.0408, 0.0338)
	(2, 0.3)	(0.0102, 0.0009)	(0.0110, 0.0010)	(0.0132, 0.0011)	(0.0192, 0.0025)
	(2, 0.4)	(0.0084, 0.0008)	(0.0091, 0.0008)	(0.0116, 0.0010)	(0.0160, 0.0024)
	(2, 0.5)	(0.0164, 0.0010)	(0.0193, 0.0011)	(0.0304, 0.0016)	(0.0310, 0.0021)
	(2, 0.6)	(0.0124, 0.0005)	(0.0162, 0.0007)	(0.0332, 0.0013)	(0.0239, 0.0010)
	(2, 0.7)	(0.0202, 0.0006)	(0.0245, 0.0007)	(0.0496, 0.0011)	(0.0392, 0.0012)
	(2, 0.8)	(0.0328, 0.0004)	(0.0406, 0.0005)	(0.0850, 0.0008)	(0.0650, 0.0009)
	(2, 0.9)	(0.0706, 0.0002)	(0.1106, 0.0003)	(0.2335, 0.0006)	(0.1391, 0.0004)
	(2, 0.95)	(0.1591, 0.0001)	(0.2423, 0.0002)	(0.5047, 0.0003)	(0.3143, 0.0003)

Table 2 Simulation results when $\alpha_0 = 2$

$$= P(B_{n,m}^{k,1}) + P(B_{n,m}^{k,2}) - P(B_{n,m}^{k,1} \cup P(B_{n,m}^{k,2})) + \cdots + P(B_{n,m}^{k,m-1}) - P\left(\left(\bigcap_{j=1}^{m-2} B_{n,m}^{k,j}\right) \cup B_{n,m}^{k,m-1}\right) + P(B_{n,m}^{k,m}) - P\left(\left(\bigcap_{j=1}^{m-1} B_{n,m}^{k,j}\right) \cup B_{n,m}^{k,m}\right).$$
(4.8)

For all $2 \le i \le m$, (4.4) implies that

= · · ·

$$\lim_{n\to\infty} P\left(\left(\bigcap_{j=1}^{i-1} B_{n,m}^{k,j}\right) \cup B_{n,m}^{k,i}\right) = 1.$$

Thus, again by (4.4), (4.7) can be proved.

Finally we prove (4.2).

Note that

$$\frac{1}{n} \max_{1 \le t \le n} g_{tk}^2(\theta_0) \le \frac{1}{n} \sup_{s \in [0,1]} \sum_{t=[ns]+1}^{[n(s+\frac{1}{m})]} g_{tk}^2(\theta_0).$$

For given $s \in [0,1]$ choose $j \in \{1,...,m\}$ so that $s \in [\frac{j-1}{m}, \frac{j}{m}]$. Then, for each $s \in [0,1]$, $\omega \in B_{n,m}^k$ implies

$$\frac{1}{n} \sum_{t=[ns]+1}^{[n(s+\frac{1}{m})]} g_{tk}^{2}(\theta_{0}) \leq \frac{1}{n} \sum_{t=[n(j-1)/m]+1}^{[n(j+1)/m]} g_{tk}^{2}(\theta_{0})$$

$$= \left(\frac{1}{n} \sum_{t=1}^{[n(j+1)/m]} g_{tk}^{2}(\theta_{0}) - \frac{j+1}{m} \sigma_{kk}\right)$$

$$- \left(\frac{1}{n} \sum_{t=1}^{[n(j-1)/m]} g_{tk}^{2}(\theta_{0}) - \frac{j-1}{m} \sigma_{kk}\right) + \frac{2}{m} \sigma_{kk}$$

$$\leq \frac{2}{m} + \frac{2}{m} \sigma_{kk}$$

$$= (1 + \sigma_{kk}) \frac{2}{m}.$$
(4.9)

Therefore, for all $m \ge 1$,

$$\lim_{n\to\infty} P\left\{\frac{1}{n}\max_{1\le t\le n}g_{tk}^2(\theta_0)\le \frac{2}{m}(1+\sigma_{kk})\right\}\ge \lim_{n\to\infty} P(B_{n,m}^k)=1,$$

showing (4.2). The proof of Lemma 4.1 is thus completed.

Lemma 4.2 Assume that (A_1) and (A_2) hold. Then

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} \left(Z_{t}^{\tau} \left(X_{t} - Z_{t}^{\tau} \alpha^{0} \right), g_{t}^{\tau} (\theta_{0}) \right)^{\tau} \stackrel{d}{\longrightarrow} N(0, M),$$

where

$$M = \begin{pmatrix} \Lambda & \Lambda_{12} \\ \Lambda_{12}^{\mathsf{r}} & \Sigma \end{pmatrix}.$$
(4.10)

Proof By the Cramer-Wold device, it suffices to show that, for all $c \in \mathbb{R}^{(p+1+r)} \setminus (0, ..., 0)$,

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}c^{\tau}\left(Z_{t}^{\tau}\left(X_{t}-Z_{t}^{\tau}\alpha^{0}\right),g_{t}^{\tau}\left(\theta_{0}\right)\right)^{\tau}\xrightarrow{d}N\left(0,c^{\tau}Mc\right).$$

For simplicity, we write $c^{\tau}(Z_t^{\tau}(X_t - Z_t^{\tau}\alpha^0), g_t^{\tau}(\theta_0))^{\tau}$ for $G_{t,c}(\theta_0)$. Further, let $\xi_{nt} = \frac{1}{\sqrt{n}}G_{t,c}(\theta_0)$ and $\mathscr{F}_{nt} = \sigma$ ($\xi_{nr}, 1 \le r \le t$). Then $\{\sum_{t=1}^{n} \xi_{nt}, \mathscr{F}_{nt}, 1 \le t \le n, n \ge 1\}$ is a zero-mean, square integrable martingale array. By making use of a martingale central limit theorem [18], it suffices to show that

$$\max_{1 \le t \le n} |\xi_{nt}| \xrightarrow{p} 0, \tag{4.11}$$

$$\sum_{t=1}^{n} \xi_{nt}^{2} \xrightarrow{p} c^{\tau} Mc, \qquad (4.12)$$

$$E\left(\max_{1 \le t \le n} \xi_{nt}^2\right) \text{ is bounded in } n, \tag{4.13}$$

and the fields are nested:

$$\mathscr{F}_{nt} \subseteq \mathscr{F}_{(n+1)t} \quad \text{for } 1 \le t \le n, n \ge 1.$$

$$(4.14)$$

Note that (4.14) is obvious. In what follows, we first consider (4.11). By a simple calculation, we have, for all $\varepsilon > 0$,

$$P\left\{\max_{1\leq t\leq n} |\xi_{nt}| > \varepsilon\right\}$$

$$\leq \sum_{t=1}^{n} P\{|\xi_{nt}| > \varepsilon\}$$

$$= \sum_{t=1}^{n} P\left\{\left|\frac{1}{\sqrt{n}}G_{t,c}(\theta_{0})\right| > \varepsilon\right\}$$

$$= nP\{|G_{t,c}(\theta_{0})| > \sqrt{n}\varepsilon\}$$

$$= n\int_{\Omega} I(|G_{t,c}(\theta_{0})| > \sqrt{n}\varepsilon) dP$$

$$\leq n\int_{\Omega} I(|G_{t,c}(\theta_{0})| > \sqrt{n}\varepsilon) \frac{(G_{t,c}(\theta_{0}))^{2}}{(\sqrt{n}\varepsilon)^{2}} dP$$

$$= \frac{1}{\varepsilon^{2}}\int_{\Omega} I(|G_{t,c}(\theta_{0})| > \sqrt{n}\varepsilon) (G_{t,c}(\theta_{0}))^{2} dP. \qquad (4.15)$$

Now by the Lebesgue control convergence theorem, we immediately find that (4.15) converges to 0 as $n \to \infty$. This settles (4.11).

Next consider (4.12). By the ergodic theorem, we have

$$\sum_{t=1}^{n} \xi_{nt}^{2} = \sum_{t=1}^{n} \left(\frac{1}{\sqrt{n}} G_{t,c}(\theta_{0}) \right)^{2}$$
$$\xrightarrow{\text{a.s.}} E \left(G_{t,c}(\theta_{0}) \right)^{2}$$
$$= c^{\tau} M c.$$

Hence (4.12) is proved.

Finally, consider (4.13). Note that $\{(\frac{1}{\sqrt{n}}G_{t,c}(\theta_0))^2, t \ge 1\}$ is a stationary sequence. Then we have

$$E\left(\max_{1\leq t\leq n}\xi_{nt}^{2}\right) = E\left(\max_{1\leq t\leq n}\left(\frac{1}{\sqrt{n}}G_{t,c}(\theta_{0})\right)^{2}\right)$$
$$\leq \frac{1}{n}E\left(\sum_{t=1}^{n}\left(G_{t,c}(\theta_{0})\right)^{2}\right)$$
$$= \frac{1}{n}\sum_{t=1}^{n}E\left(G_{t,c}(\theta_{0})\right)^{2}$$
$$= c^{\mathrm{T}}Mc.$$
(4.16)

This proves that (4.13). Thus, we complete the proof of Lemma 4.2.

Lemma 4.3 Assume that (A_1) and (A_2) hold. Then

$$\lambda_{\theta_0} = O_p(n^{-\frac{1}{2}}).$$

Proof Let $\lambda_{\theta_0} = \zeta \beta_0$, where $\|\beta_0\| = 1$ is a unit vector and $\zeta = \|\lambda_{\theta_0}\|$. Then (2.3) implies that

$$\begin{aligned} 0 &= \frac{\beta_0^{\tau}}{n} \sum_{t=1}^n \frac{g_t(\theta_0)}{1 + \zeta \beta_0^{\tau} g_t(\theta_0)} \\ &= \frac{\beta_0^{\tau}}{n} \sum_{t=1}^n g_t(\theta_0) - \frac{\zeta}{n} \sum_{t=1}^n \frac{(\beta_0^{\tau} g_t(\theta_0))^2}{1 + \zeta \beta_0^{\tau} g_t(\theta_0)} \\ &\leq \frac{\beta_0^{\tau}}{n} \sum_{t=1}^n g_t(\theta_0) - \frac{\zeta}{1 + \zeta \max_{1 \le t \le n} \|g_t(\theta_0)\|} \beta_0^{\tau} \Sigma_n(\theta_0) \beta_0. \end{aligned}$$

This implies that

$$\zeta \beta_0^{\tau} \Sigma_n(\theta) \beta_0 - \max_{1 \le t \le n} \left\| g_t(\theta_0) \right\| \frac{\beta_0^{\tau}}{n} \sum_{t=1}^n g_t^{\tau}(\theta_0) \le \frac{\beta_0^{\tau}}{n} \sum_{t=1}^n g_t^{\tau}(\theta_0).$$
(4.17)

Note that

$$\left|\frac{\beta_0^{\tau}}{n}\sum_{t=1}^n g_t^{\tau}(\theta_0)\right| \le \left\|\frac{1}{n}\sum_{t=1}^n g_t^{\tau}(\theta_0)\right\| = O_p(n^{-\frac{1}{2}}).$$
(4.18)

By using Lemma 4.1 and (4.18), we can obtain

$$\max_{1 \le t \le n} \left\| g_t(\theta_0) \right\| \frac{\beta_0^{\tau}}{n} \sum_{t=1}^n g_t^{\tau}(\theta_0) = o_p(1).$$
(4.19)

By (4.3), we have

$$\beta_0^{\tau} \Sigma_n(\theta_0) \beta_0 \xrightarrow{p} \beta_0^{\tau} \Sigma(\theta_0) \beta_0.$$
(4.20)

By this, together with (4.17)-(4.19), we can prove Lemma 4.3. $\hfill \Box$

Proof of Theorem 2.1 By (2.3), we have

$$\lambda_{\theta_0} = \left(\Sigma_n(\theta_0)\right)^{-1} \frac{1}{n} \sum_{t=1}^n g_t(\theta_0) + \left(\Sigma_n(\theta_0)\right)^{-1} R_n(\theta_0),$$
(4.21)

where

$$R_n(\theta_0) = \frac{1}{n} \sum_{t=1}^n g_t^{\tau}(\theta_0) \frac{(\lambda_{\theta_0}^{\tau} g_t(\theta_0))^2}{1 + \lambda_{\theta_0}^{\tau} g_t(\theta_0)}.$$

By Lemmas 4.1-4.3, we know that

$$R_n(\theta_0) = o_p(n^{-\frac{1}{2}}).$$
(4.22)

This implies uniformly for t

$$\omega_t(\theta_0) = \frac{1}{n} \frac{1}{1 + \lambda_{\theta_0}^{\tau} g_t(\theta_0)}$$

= $\frac{1}{n} \left(1 - \lambda_{\theta_0}^{\tau} g_t(\theta_0) (1 + o_p(1)) \right).$ (4.23)

Moreover, note that

$$\sqrt{n}(\hat{\alpha} - \alpha^0) = T_n^{-1} S_n, \tag{4.24}$$

where $T_n = \sum_{t=1}^n \omega_t(\theta_0) Z_t Z_t^{\tau}$ and $S_n = \sqrt{n} \sum_{t=1}^n \omega_t(\theta_0) Z_t (X_t - Z_T^{\tau} \alpha^0)$. First, we consider T_n . By (4.23), we have

$$\begin{split} T_n &= \sum_{t=1}^n \omega_t(\theta_0) Z_t Z_t^{\tau} \\ &= \frac{1}{n} \sum_{t=1}^n \left(1 - \lambda_{\theta_0}^{\tau} g_t(\theta_0) \left(1 + o_p(1) \right) \right) Z_t Z_t^{\tau} \\ &= \frac{1}{n} \sum_{t=1}^n Z_t Z_t^{\tau} - \frac{1}{n} \sum_{t=1}^n (Z_t Z_t^{\tau}) \otimes \left(\left(\left(\Sigma_n(\theta_0) \right)^{-1} R_n(\theta_0) \right)^{\tau} g_t(\theta_0) \left(1 + o_p(1) \right) \right) \\ &- \frac{1}{n} \sum_{t=1}^n (Z_t Z_t^{\tau}) \otimes \left(\left(\left(\Sigma_n(\theta_0) \right)^{-1} \frac{1}{n} \sum_{t=1}^n g_t(\theta_0) \right)^{\tau} g_t(\theta_0) \left(1 + o_p(1) \right) \right) \\ &\doteq U_1 - U_2 - U_3. \end{split}$$

Then by (4.3) and (4.22), conditions (A_1) and (A_2) , Lemma 4.2, and the ergodic theorem, we can prove that

$$U_2 = o_p(1). (4.25)$$

Similarly, we can obtain

$$U_3 \xrightarrow{\text{a.s.}} 0.$$
 (4.26)

This, together with (4.25), yields

$$T_n \xrightarrow{p} W.$$
 (4.27)

Next consider S_n . By (4.23), we have

$$\begin{split} S_n &= \sqrt{n} \sum_{t=1}^n \omega_t(\theta_0) Z_t \big(X_t - Z_t^{\mathsf{T}} \alpha^0 \big) \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \big(1 - \lambda_{\theta_0}^{\mathsf{T}} g_t(\theta_0) \big(1 + o_p(1) \big) \big) Z_t \big(X_t - Z_t^{\mathsf{T}} \alpha^0 \big). \end{split}$$

Thus, combining with (4.21), (4.22), we can obtain

$$S_{n} = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} Z_{t} \left(X_{t} - Z_{t}^{\tau} \alpha^{0} \right)$$
$$- \frac{1}{n} \sum_{t=1}^{n} \left(X_{t} - Z_{t}^{\tau} \alpha^{0} \right) Z_{t} g_{t}^{\tau} (\theta_{0}) \left(\Sigma_{n}(\theta_{0}) \right)^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} g_{t}(\theta_{0}) + o_{p}(1).$$

By the ergodic theorem, we have

$$\frac{1}{n}\sum_{t=1}^{n} (X_t - Z_t^{\tau}\alpha^0) Z_t g_t^{\tau}(\theta_0) \xrightarrow{\text{a.s.}} \Lambda_{12}.$$

This, together with (4.3) and Lemma 4.2, proves that

$$S_n \xrightarrow{d} N(0, \Lambda - \Lambda_{12} \Sigma^{-1} \Lambda_{12}^{\tau}),$$
 (4.28)

which, combining with (4.27), proves Theorem 2.1.

Proof of Theorem 2.2 Similar to the proof of Lemma 1 and Theorem 1 in Qin and Lawless [16], we can prove that

$$\lambda_{\hat{\theta}} = B \frac{1}{n} \sum_{t=1}^{n} g_t(\theta_0) + o_p(n^{-\frac{1}{2}}).$$
(4.29)

Moreover, note that

$$\sqrt{n}\left(\hat{\alpha}^{1}-\alpha^{0}\right)=\tilde{T}_{n}^{-1}\tilde{S}_{n},\tag{4.30}$$

where

$$\tilde{T}_n = \sum_{t=1}^n \omega_t(\hat{\theta}) Z_t Z_t^{\tau}$$

and

$$\tilde{S}_n = \sqrt{n} \sum_{t=1}^n \omega_t(\hat{\theta}) Z_t \big(X_t - Z_t^{\tau} \alpha^0 \big).$$

By an argument similar to the proof of Theorem 2.1, we can prove that

$$\tilde{T}_n \xrightarrow{p} W$$
 (4.31)

and

$$\tilde{S}_n \xrightarrow{d} N(0, \Lambda - \Lambda_{12} B \Lambda_{12}^{\tau}),$$
(4.32)

showing (2.7). The proof of Theorem 2.2 is thus completed.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

The authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

Author details

¹ College of Mathematics, Jilin University, Changchun, 130012, China. ²Public Foreign Languages Department, Jilin Normal University, Siping, 136000, China. ³ College of Mathematics, Jilin Normal University, Siping, 136000, China.

Acknowledgements

This work is supported by National Natural Science Foundation of China (Nos. 11271155, 11001105, 11071126, 10926156, 11071269), Specialized Research Fund for the Doctoral Program of Higher Education (Nos. 20110061110003, 20090061120037), Scientific Research Fund of Jilin University (Nos. 201100011, 200903278), the Science and Technology Development Program of Jilin Province (201201082), Jilin Province Social Science Fund (2012B115), and Jilin Province Natural Science Foundation (20101596, 20130101066JC).

Received: 8 November 2014 Accepted: 2 June 2015 Published online: 02 July 2015

References

- 1. Alzaid, AA, Al-Osh, M: An integer-valued pth-order autoregressive structure (INAR(p)) process. J. Appl. Probab. 27, 314-324 (1990)
- Davis, RA, Dunsmuir, WTM, Streett, SB: Observation-driven models for Poisson counts. Biometrika 90, 777-790 (2003)
 Zheng, H, Basawa, IV, Datta, S: Inference for *p*th-order random coefficient integer-valued autoregressive processes.
- J. Time Ser. Anal. **27**, 411-440 (2006)
- 4. Weiß, CH: Thinning operations for modeling time series of counts a survey. AStA Adv. Stat. Anal. 92, 319-341 (2008)
- 5. Ferland, R, Latour, A, Oraichi, D: Integer-valued GARCH process. J. Time Ser. Anal. 27, 923-942 (2006)
- 6. Zhu, F, Wang, D: Estimation and testing for a Poisson autoregressive model. Metrika 73, 211-230 (2011)
- 7. Hansen, LP: Large sample properties of generalized method of moments estimators. Econometrica 50, 1029-1054 (1982)
- 8. Isaki, CT: Variance estimation using auxiliary information. J. Am. Stat. Assoc. 78, 117-123 (1983)
- 9. Kuk, AYC, Mak, TK: Median estimation in the presence of auxiliary information. J. R. Stat. Soc. B 51, 261-269 (1989)
- Rao, JNK, Kovar, JG, Mantel, HJ: On estimating distribution functions and quantiles from survey data using auxiliary information. Biometrika 77, 365-375 (1990)
- 11. Owen, AB: Empirical likelihood ratio confidence intervals for a single functional. Biometrika 75, 237-249 (1988)
- 12. Owen, AB: Empirical likelihood and confidence region. Ann. Stat. 18, 90-120 (1990)
- Chen, J, Qin, J: Empirical likelihood estimation for finite populations and the effective usage of auxiliary information. Biometrika 80, 107-116 (1993)
- Zhang, B: M-Estimation and quantile estimation in the presence of auxiliary information. J. Stat. Plan. Inference 44, 77-94 (1995)
- Tang, CY, Leng, C: An empirical likelihood approach to quantile regression with auxiliary information. Stat. Probab. Lett. 82, 29-36 (2012)
- 16. Qin, J, Lawless, J: Empirical likelihood and general estimating equations. Ann. Stat. 22, 300-325 (1994)
- 17. Owen, AB: Empirical Likelihood. Chapman & Hall/CRC, London (2001)
- 18. Hall, P, Heyde, CC: Martingale Limit Theory and Its Application. Academic Press, New York (1980)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ► Convenient online submission
- ► Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at > springeropen.com