Journal of Mathematical Research & Exposition
Sept., 2009, Vol. 29, No. 5, pp. 923–930
DOI:10.3770/j.issn:1000-341X.2009.05.020
Http://jmre.dlut.edu.cn

Empirical Likelihood Inference for MA(q) Model

CHEN Yan Hong, SONG Li Xin
(School of Mathematical Sciences, Dalian University of Technology, Liaoning 116024, China)
(E-mail: cyh_0825@163.com)

Abstract In this article we study the empirical likelihood inference for MA(q) model. We propose the moment restrictions, by which we get the empirical likelihood estimator of the model parameter, and we also propose an empirical log-likelihood ratio based on this estimator. Our result shows that the EL estimator is asymptotically normal, and the empirical log-likelihood ratio is proved to be asymptotical standard chi-square distribution.

Keywords MA(q); empirical likelihood; moment restriction; asymptotic properties.

Document code A
MR(2000) Subject Classification 62M10; 62G10; 62E20
Chinese Library Classification O212.7

1. Introduction

In this paper, we consider such time series. Suppose $\{X_t\}$ satisfy the following time series model:

$$X_t = \varepsilon_t + \beta_1\varepsilon_{t-1} + \beta_2\varepsilon_{t-2} + \cdots + \beta_q\varepsilon_{t-q}, \tag{1}$$

where $\{\varepsilon_t\}$ are independent identical distribution random variables such that $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = 1$. Let $\beta = (\beta_1, \beta_2, \ldots, \beta_q)^\tau$ denote (q) dimension parameter vector, and "$^\tau$" stands for the transpose of a vector or a matrix, and we use $\beta^{(0)}$ to stand for the real parameter value.

The empirical likelihood (EL) method was first proposed by Owen (1988) and its general property was subsequently studied by Owen (1990). From then on, it has taken much of the spotlight in the statistical literature. As for time series models, Monti (1997) derived the EL confidence regions in time series models by a spectral method, and applied them in ARMA model. However, in this paper we get EL inference by constructing moment restrictions. The main objective of this paper is to apply the EL estimation to model (1). Then we derive the EL estimator from moment restrictions, and prove it is asymptotically normal. In order to examine the estimator, we propose empirical log-likelihood ratio whose asymptotic distribution is exactly a standard chi-squared.

The paper is organized as follows. In Section 2, EL estimator of the parameter in the model is derived by moment restrictions, and the test problem based on this estimator is also proposed. The asymptotic properties are investigated in Section 3.

Received date: 2007-10-18; **Accepted date**: 2008-11-09
Foundation item: "Mathematics+X" of DLUT (No. 842306).

2. Empirical likelihood method

By the moment and independencies of $\{\varepsilon_t\}$, we now propose the restriction conditions. Let $g_t(\beta) = (g_{1t}(\beta), g_{2t}(\beta), \ldots, g_{(q+1)t}(\beta))^\tau$, and

$$\begin{aligned} g_{1t}(\beta) &= X_t^2 - E(X_t^2) = X_t^2 - (1+\beta_1^2+\beta_2^2+\cdots+\beta_q^2), \\ g_{2t}(\beta) &= X_tX_{t-1} - E(X_tX_{t-1}) = X_tX_{t-1} - (\beta_1+\beta_2\beta_1+\cdots+\beta_q\beta_{q-1}), \\ &\cdots\ \cdots\ \cdots\ \cdots \\ g_{qt}(\beta) &= X_tX_{t-q} - E(X_tX_{t-q}) = X_tX_{t-q} - \beta_q. \end{aligned} \tag{2}$$

We can easily get $Eg_t(\beta^{(0)}) = 0$.

Suppose that $X_1, X_2, \ldots, X_n$ are observed samples of the model (1). Let $p_t = P(X = X_t)$ $(t = 1, 2, \ldots, n)$. We have the following moment restrictions:

$$\sum_{t=1}^{n} p_t g_t(\beta) = 0. \tag{3}$$

In order to get the EL estimator, we should consider the maximum point of the function

$$L_n(\beta) = \prod_{t=1}^{n} p_t$$

on the set of $\mathcal{D}_n(\beta) = \{\ p_t \mid p_t \geq 0, \sum_{t=1}^n p_t = 1, \sum_{t=1}^n p_t g_t(\beta) = 0\ \}$, where $g_t(\beta)$ is as in (2).

It is not difficult to prove that when $\beta \in B(\beta^{(0)}, n^{-\frac{1}{3}})$, there exist $\{p_t\}_{t=1}^n$, where $p_t > 0$, and $\sum_{t=1}^n p_t = 1$, such that $\sum_{t=1}^n p_t g_t(\beta) = 0$, where $B(\beta^{(0)}, n^{-\frac{1}{3}})$ is defined to be the set $\{\beta : \|\beta - \beta^{(0)}\| \leq n^{-\frac{1}{3}}\}$. We can easily prove that $L_n(\beta)$ is a convex function, and for given β, $\mathcal{D}_n(\beta)$ is a bounded convex set, so the extremum point on the set of $\mathcal{D}_n(\beta)$ is also the maximum point. Since $\sup_{p_t} \prod_{t=1}^n p_t > 0$, we can further prove that the maximum point is unique. So we can use Lagrange multiplier technique and Kuhn–Tucker condition to obtain the optimal $\{p_t\}_{t=1}^n$, just as Owen (1991) did in linear models. Let

$$G = \sum_{t=1}^{n} \ln p_t + \mu(1 - \sum_{t=1}^{n} p_t) - n\lambda^\tau \sum_{t=1}^{n} p_t g_t(\beta) + \sum_{t=1}^{n} \nu_t p_t,$$

where μ, λ, ν_t $(t = 1, 2, \ldots, n)$ are Lagrange multipliers. Differentiating G with respect to p_t, we get

$$\frac{\partial G}{p_t} = \frac{1}{p_t} - \mu - n\lambda^\tau g_t(\beta) + \nu_t = 0, \quad \nu_t p_t = 0, \quad t = 1, 2, \ldots, n$$

and

$$p_t = \frac{1}{n\,[1 + \lambda^\tau g_t(\beta)]}$$

subject to the last restriction in $\mathcal{D}_n(\beta)$:

$$\sum_{t=1}^{n} p_t g_t(\beta) = \frac{1}{n} \sum_{t=1}^{n} \frac{g_t(\beta)}{1 + \lambda^\tau g_t(\beta)} = 0. \tag{4}$$

Notice that,

$$\frac{\partial}{\lambda}\Big\{\frac{1}{n}\sum_{t=1}^{n} \frac{1}{1+\lambda^\tau g_t(\beta)} g_t(\beta)\Big\} = -\frac{1}{n}\sum_{t=1}^{n} \frac{g_t(\beta) g_t^\tau(\beta)}{[1+\lambda^\tau g_t(\beta)]^2}.$$

By implicit function theorem, λ can be regarded as a continuous differentiable function of β, and is relative to n, written as $\lambda_n(\beta)$. For given β, the log EL function can be defined as

$$l_n(\beta) = \sum_{t=1}^{n} \log[1 + \lambda_n^\tau(\beta) g_t(\beta)] \tag{5}$$

whose minimum is just the maximum of $L_n(\beta)$. So $\widehat{\beta}_n = \arg\min l_n(\beta)$, and the corresponding Lagrange multiplier can be noted as $\widehat{\lambda}_n \triangleq \lambda_n(\widehat{\beta}_n)$.

In order to indicate the validity of our estimator, we should consider the test problem, that is, test null hypothesis $H_0 : \beta = \beta^{(0)}$. We can define the empirical log-likelihood ratio as

$$T = 2l_n(\beta^{(0)}) - 2l_n(\widehat{\beta}_n). \tag{6}$$

In the next section, we mainly discuss the asymptotic properties of our EL estimator and empirical log-likelihood ratio.

3. Asymptotic properties

In order to gain distinct results, we require three additional conditions.

Assumption 1 $E|X_t|^6 < \infty$.

Assumption 2 The rank of matrix $E\frac{\partial g_t(\beta)}{\partial \beta}|_{\beta=\beta^{(0)}}$ is q.

Assumption 3 Matrix C is positive definite, which is defined as:

$$C \triangleq E(g_t(\beta^{(0)}) g_t^\tau(\beta^{(0)})) + \sum_{j=1}^{2q} E(g_t(\beta^{(0)}) g_{t-j}^\tau(\beta^{(0)})) + \sum_{j=1}^{2q} E(g_{t-j}(\beta^{(0)}) g_t^\tau(\beta^{(0)})).$$

Owing to special cases, we know that all three assumptions above are reasonable.

By Chen (1997) we can easily prove that $\frac{1}{n}\sum_{t=1}^{n} g_t(\beta^{(0)}) = O(n^{-\frac{1}{2}} \log\log n)$ almost surely. And when we suppose Assumption 3 holds, it is also easy to prove that the distribution of $\frac{1}{\sqrt{n}}\sum_{t=1}^{n} g_t(\beta^{(0)})$ is asymptotic normal, that is

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} g_t(\beta^{(0)}) \xrightarrow{L} N(0, C).$$

Before giving the main results, we need some conclusions as in the following lemma, which are required in the proof of the main results.

Lemma 1 *Under the assumptions above, when $\beta \in B(\beta^{(0)}, n^{-\frac{1}{3}})$ (in the following proof it is denoted as B), and for n large enough, the following conclusions hold:*

1) (i) $\sup_{\beta \in B} E||g_t(\beta)||^3 = O(1)$,

(ii) $\sup_{\beta \in B} E||\frac{\partial g_t(\beta)}{\partial \beta}|| = O(1)$ *and* $\sup_{\beta \in B} E||\frac{\partial^2 g_t(\beta)}{\partial \beta \partial \beta^\tau}|| = O(1)$.

2) $\sup_{\beta \in B} \|\frac{1}{n}\sum_{t=1}^{n} g_t(\beta)|| = O_p(n^{-\frac{1}{3}})$.

3) (i) $\sup_{\beta \in B} \|\frac{1}{n}\sum_{t=1}^{n} g_t(\beta) g_t^\tau(\beta) - E g_t(\beta^{(0)}) g_t^\tau(\beta^{(0)})\| = o_p(1)$, *and* $E[g_t(\beta^{(0)}) g_t^\tau(\beta^{(0)})]$ *is a positive definite matrix.*

(ii) $\sup_{\beta\in B}\|[\frac{1}{n}\sum_{t=1}^{n} g_t(\beta)g_t^{\tau}(\beta)]^{-1}-[Eg_t(\beta^{(0)})g_t^{\tau}(\beta^{(0)})]^{-1}\| = o_p(1)$, *and* $[Eg_t(\beta^{(0)})g_t^{\tau}(\beta^{(0)})]^{-1}$ *is a positive definite matrix.*

4) *(i)* $\sup_{\beta\in B}\|\frac{1}{n}\sum_{t=1}^{n}\frac{\partial g_t(\beta)}{\partial\beta} - E\frac{\partial g_t(\beta^{(0)})}{\partial\beta}\| = o(1)$.

(ii) $\sup_{\beta\in B}\|\frac{1}{n}\sum_{t=1}^{n}\frac{\partial^2 g_t(\beta)}{\partial\beta\partial\beta^{\tau}} - E\frac{\partial^2 g_t(\beta^{(0)})}{\partial\beta\partial\beta^{\tau}}\| = o(1)$.

5) *Let* $z_n = \max_{1\le t\le n}\sup_{\beta\in B}\|g_t(\beta)\|$. *Then* $z_n = o(n^{\frac{1}{3}})$.

6) $\sup_{\beta\in B}\|\lambda_n(\beta)\| = O_p(n^{-\frac{1}{3}})$.

7) $\sup_{\beta\in B}\|\lambda_n(\beta) - [\frac{1}{n}\sum_t g_t(\beta)g_t^{\tau}(\beta)]^{-1}\cdot[\frac{1}{n}\sum_t g_t(\beta)]\| = o_p(n^{-\frac{1}{3}})$.

Theorem 2 *Under the assumptions above, when* $\beta\in B=\{\beta:\|\beta-\beta^{(0)}\|\le n^{-\frac{1}{3}}\}$, *we have*

$$P(\widehat{\beta}_n \text{ is the interior point of } B)\to 1, \text{as } n\to\infty.$$

Furthermore when $\widehat{\beta}_n$ *is the intetior point, then*

$$Q_{1n}(\widehat{\beta}_n,\widehat{\lambda}_n) = \frac{1}{n}\sum_t \frac{g_t(\beta)}{1+\lambda_n^{\tau}(\beta)g_t(\beta)}\Big|_{\beta=\widehat{\beta}_n} = 0,$$

$$Q_{2n}(\widehat{\beta}_n,\widehat{\lambda}_n) = \frac{1}{n}\sum_t \frac{1}{1+\lambda_n^{\tau}(\beta)g_t(\beta)}(\frac{\partial g_t(\beta)}{\partial\beta})^{\tau}\lambda_n(\beta)\Big|_{\beta=\widehat{\beta}_n} = 0. \tag{7}$$

Proof When $\beta\in W=\{\beta|\ \|\beta-\beta^{(0)}\|=n^{-\frac{1}{3}}\}$, denoted as $\beta=\beta^{(0)}+\mu\cdot n^{-\frac{1}{3}}$, then by the conclusions of the lemma above, we can easily have

$$\frac{1}{n}\sum_t g_t(\beta) = \frac{1}{n}\sum_t g_t(\beta^{(0)}) + \frac{1}{n}\sum_t \frac{\partial g_t(\beta^{(0)})}{\partial\beta}\mu n^{-\frac{1}{3}} + O_p(n^{-\frac{2}{3}})$$

and

$$l_n(\beta) = \frac{n}{2}\,[\frac{1}{n}\sum_t g_t(\beta)]^{\tau}\cdot[\frac{1}{n}\sum_t g_t(\beta)g_t^{\tau}(\beta)]^{-1}\cdot[\frac{1}{n}\sum_t g_t(\beta)] + o_p(n^{\frac{1}{3}}).$$

So

$$l_n(\beta) = \frac{1}{2}\cdot n^{\frac{1}{3}}\mu^{\tau}(E(\frac{\partial g_t(\beta^{(0)})}{\partial\beta}))^{\tau}[Eg_t(\beta^{(0)})g_t^{\tau}(\beta^{(0)})]^{-1}E(\frac{\partial g_t(\beta^{(0)})}{\partial\beta})\mu + o_p(n^{\frac{1}{3}}).$$

Because matrix $\Lambda = (E(\frac{\partial g_t(\beta^{(0)})}{\partial\beta}))^{\tau}[Eg_t(\beta^{(0)})g_t^{\tau}(\beta^{(0)})]^{-1}E(\frac{\partial g_t(\beta^{(0)})}{\partial\beta})$ is a positive definite matrix, such that $\mu^{\tau}\Lambda\mu > c_0$, where $c_0>0$ is the minimum latent root of Λ. And then we get

$$\inf_{\beta\in W} l_n(\beta) \ge \frac{1}{2}c_0\cdot n^{\frac{1}{3}} + o_p(n^{\frac{1}{3}}). \tag{8}$$

When $\beta=\beta^{(0)}$, we know $\frac{1}{n}\sum_t g_t(\beta^{(0)}) = O(n^{-\frac{1}{2}}\log\log n)$. Performing the similar proof for $l_n(\beta)$, we can obtain

$$\begin{aligned} l_n(\beta^{(0)}) &= \frac{n}{2}[\frac{1}{n}\sum_t g_t(\beta^{(0)})]^{\tau}[\frac{1}{n}\sum_t g_t(\beta^{(0)})g_t^{\tau}(\beta^{(0)})]^{-1}[\frac{1}{n}\sum_t g_t(\beta^{(0)})] + o_p(1) \\ &= O_p\ (\log\log n). \end{aligned} \tag{9}$$

Because $l_n(\beta)$ is a continuous function about β, it follows from (8) and (9) that

$$\begin{aligned} P(\ \widehat{\beta}_n \text{ is the interior point of B}) &\ge P\Big(\frac{\inf_{\beta\in W}\ln(\beta)-\ln(\beta^{(0)})}{n^{\frac{1}{3}}} \ge \frac{1}{4}\cdot c_0\Big) \\ &\ge P\Big(\frac{\frac{1}{2}\cdot c_0\cdot n^{\frac{1}{3}} + o_p(n^{\frac{1}{3}}) - O_p(\log\log n)}{n^{\frac{1}{3}}} \ge \frac{1}{4}\cdot c_0\Big) \to 1, \quad n\to\infty. \end{aligned}$$

So the minimum value point $\widehat{\beta}_n$ of $l_n(\beta)$ is interior point which satisfies

$$0 = \frac{\partial l_n(\beta)}{\partial \beta}\Big|_{\beta=\widehat{\beta}_n} = \sum_t \frac{\frac{\partial \lambda_n^\tau(\beta)}{\partial \beta} g_t(\beta) + \frac{\partial g_t^\tau(\beta)}{\partial \beta}\lambda_n(\beta)}{1+\lambda_n^\tau(\beta) g_t(\beta)}\Big|_{\beta=\widehat{\beta}_n}$$
$$= (\frac{\partial \lambda_n(\beta)}{\partial \beta})^\tau \sum_t \frac{g_t(\beta)}{1+\lambda_n^\tau(\beta) g_t(\beta)}\Big|_{\beta=\widehat{\beta}_n} + \sum_t \frac{1}{1+\lambda_n^\tau(\beta) g_t(\beta)} (\frac{\partial g_t(\beta)}{\partial \beta})^\tau \lambda_n(\beta)\Big|_{\beta=\widehat{\beta}_n}.$$

Therefore,

$$Q_{1n}(\widehat{\beta}_n, \widehat{\lambda}_n) = 0, \quad Q_{2n}(\widehat{\beta}_n, \widehat{\lambda}_n) = 0.$$

□

Theorem 3 *Under the assumptions above,*

$$\sqrt{n}(\widehat{\beta}_n - \beta^{(0)}) \xrightarrow{L} N(0, V), \quad \sqrt{n}(\widehat{\lambda}_n - 0) \xrightarrow{L} N(0, U),$$

where

$$V = (-S_{22.1}^{-1} S_{21} S_{11}^{-1}) C (-S_{22.1}^{-1} S_{21} S_{11}^{-1})^\tau,$$
$$U = (S_{11}^{-1} + S_{11}^{-1} S_{12} S_{22.1}^{-1} S_{21} S_{11}^{-1}) C (S_{11}^{-1} + S_{11}^{-1} S_{12} S_{22.1}^{-1} S_{21} S_{11}^{-1})^\tau.$$

Proof Since

$$\begin{aligned} 0 =& Q_{1n}(\widehat{\beta}_n, \widehat{\lambda}_n) = Q_{1n}(\beta^{(0)}, 0) + \frac{\partial Q_{1n}(\beta, \lambda)}{\partial \beta}\Big|_{\beta=\beta^{(0)}, \lambda=0} (\widehat{\beta}_n - \beta^{(0)}) + \\ & \frac{\partial Q_{1n}(\beta, \lambda)}{\partial \lambda^\tau}\Big|_{\beta=\beta^{(0)}, \lambda=0} (\widehat{\lambda}_n - 0) + O_p(\delta_n), \\ 0 =& Q_{2n}(\widehat{\beta}_n, \widehat{\lambda}_n) = Q_{2n}(\beta^{(0)}, 0) + \frac{\partial Q_{1n}(\beta, \lambda)}{\partial \beta}\Big|_{\beta=\beta^{(0)}, \lambda=0} (\widehat{\beta}_n - \beta^{(0)}) + \\ & \frac{\partial Q_{2n}(\beta, \lambda)}{\partial \lambda^\tau}\Big|_{\beta=\beta^{(0)}, \lambda=0} (\widehat{\lambda}_n - 0) + O_p(\delta_n) \end{aligned} \tag{10}$$

where $\delta_n = \|\widehat{\beta}_n - \beta^{(0)}\|^2 + \|\widehat{\lambda}_n\|^2 = O_p(n^{-\frac{2}{3}})$, and

$$Q_{1n}(\beta^{(0)}, 0) = \frac{1}{n}\sum_t g_t(\beta^{(0)}), \quad Q_{2n}(\beta^{(0)}, 0) = 0,$$
$$\frac{\partial Q_{1n}(\beta, \lambda)}{\partial \beta}\Big|_{\beta=\beta^{(0)}, \lambda=0} = \frac{1}{n}\sum_t \frac{\partial g_t(\beta^{(0)})}{\partial \beta},$$
$$\frac{\partial Q_{2n}(\beta, \lambda)}{\partial \beta}\Big|_{\beta=\beta^{(0)}, \lambda=0} = 0,$$
$$\frac{\partial Q_{1n}(\beta, \lambda)}{\partial \lambda^\tau}\Big|_{\beta=\beta^{(0)}, \lambda=0} = -\frac{1}{n}\sum_t g_t(\beta^{(0)}) g_t^\tau(\beta^{(0)}),$$
$$\frac{\partial Q_{2n}(\beta, \lambda)}{\partial \lambda^\tau}\Big|_{\beta=\beta^{(0)}, \lambda=0} = \frac{1}{n}\sum_t (\frac{\partial g_t(\beta^{(0)})}{\partial \beta})^\tau,$$

by (10), we get

$$\begin{pmatrix} \widehat{\lambda}_n - 0 \\ \widehat{\beta}_n - \beta^{(0)} \end{pmatrix} = S_n^{-1} \begin{pmatrix} -Q_{1n}(\beta^{(0)},\ 0) + o_p(n^{-\frac{1}{2}}) \\ o_p(n^{-\frac{1}{2}}) \end{pmatrix}$$

where

$$S_n = \begin{pmatrix} -\frac{1}{n}\sum_t g_t(\beta^{(0)})g_t^\tau(\beta^{(0)}) & \frac{1}{n}\sum_t \frac{\partial g_t(\beta^{(0)})}{\partial \beta} \\ \frac{1}{n}\sum_t(\frac{\partial g_t(\beta^{(0)})}{\partial \beta})^\tau & 0 \end{pmatrix} \xrightarrow{\text{a.s.}}$$

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} = \begin{pmatrix} -E[g_t(\beta^{(0)})g_t^\tau(\beta^{(0)})] & E(\frac{\partial g_t(\beta^{(0)})}{\partial \beta}) \\ E(\frac{\partial g_t(\beta^{(0)})}{\partial \beta})^\tau & 0 \end{pmatrix}$$

and

$$S_n^{-1} \to S^{-1} = \begin{pmatrix} S_{11}^{-1} + S_{11}^{-1}S_{12}S_{22.1}^{-1}S_{21}S_{11}^{-1} & -S_{11}^{-1}S_{12}S_{22.1}^{-1} \\ -S_{22.1}^{-1}S_{21}S_{11}^{-1} & S_{22.1}^{-1} \end{pmatrix} \triangleq \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix},$$

where $S_{22.1} = S_{22} - S_{21}S_{11}^{-1}S_{12} = -S_{21}S_{11}^{-1}S_{12}$. Therefore

$$\sqrt{n}\begin{pmatrix} \widehat{\lambda}_n - 0 \\ \widehat{\beta}_n - \beta^{(0)} \end{pmatrix} = S_n^{-1}\begin{pmatrix} -\sqrt{n}Q_{1n}(\beta^{(0)},\ 0) + o_p(1) \\ o_p(1) \end{pmatrix}.$$

Let $n \to \infty$. By the asymptotic normality of $\sqrt{n}Q_{1n}(\beta^{(0)}, 0)$, we get

$$\sqrt{n}\begin{pmatrix} \widehat{\lambda}_n - 0 \\ \widehat{\beta}_n - \beta^{(0)} \end{pmatrix} \xrightarrow{L} N(0, W),$$

and the relevant covariance matrix W is equal to

$$\begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \cdot \begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} T_{11}^\tau & T_{21}^\tau \\ T_{12}^\tau & T_{22}^\tau \end{pmatrix} = \begin{pmatrix} T_{11}CT_{11}^\tau & T_{11}CT_{21}^\tau \\ T_{21}CT_{11}^\tau & T_{21}CT_{21}^\tau \end{pmatrix},$$

where

$$C = Eg_t(\beta^{(0)})g_t^\tau(\beta^{(0)}) + \sum_{j=1}^{2q} Eg_t(\beta^{(0)})g_{t-j}^\tau(\beta^{(0)}) + \sum_{j=1}^{2q} Eg_{t-j}(\beta^{(0)})g_t^\tau(\beta^{(0)}).$$

And we define V, U respectively as

$$V \triangleq T_{21}CT_{21}^\tau = (-S_{22.1}^{-1}S_{21}S_{11}^{-1})C(-S_{22.1}^{-1}S_{21}S_{11}^{-1})^\tau,$$
$$U \triangleq T_{11}CT_{11}^\tau = (S_{11}^{-1} + S_{11}^{-1}S_{12}S_{22.1}^{-1}S_{21}S_{11}^{-1})C(S_{11}^{-1} + S_{11}^{-1}S_{12}S_{22.1}^{-1}S_{21}S_{11}^{-1})^\tau,$$

especially,

$$\sqrt{n}(\widehat{\beta}_n - \beta^{(0)}) = S_{22.1}^{-1}S_{21}S_{11}^{-1}\sqrt{n}Q_{1n}(\beta^{(0)}, 0) + o_p(1),$$

and

$$\sqrt{n}(\widehat{\lambda}_n - 0) = -(S_{11}^{-1} + S_{11}^{-1}S_{12}S_{22.1}^{-1}S_{21}S_{11}^{-1})\sqrt{n}Q_{1n}(\beta^{(0)}, 0) + o_p(1).$$

We get

$$\sqrt{n}(\widehat{\beta}_n - \beta^{(0)}) \xrightarrow{L} N(0, V), \quad \sqrt{n}(\widehat{\lambda}_n - 0) \xrightarrow{L} N(0, U).$$

Moreover, the EL estimators $\widehat{\beta}_n$ and $\widehat{\lambda}_n$ are asymptotically irrelative. □

The theorem above shows the congruence consistency and asymptotic normality of the EL estimator, while the following theorem reveals that the distribution of the empirical log-likelihood ratio (6) is exactly a standard chi-squared.

Theorem 4 *If $H_0 : \beta = \beta^{(0)}$, then under the assumptions above, we have*

$$T \xrightarrow{L} \chi_q^2,$$

where $T = 2l_n(\beta^{(0)}) - 2l_n(\widehat{\beta}_n)$.

Proof According to the asymptotical normality of EL estimator, we have $\widehat{\beta}_n - \beta^{(0)} = O_p(n^{-\frac{1}{2}})$, $\widehat{\lambda}_n = O_p(n^{-\frac{1}{2}})$. Using the above analogous method, we can prove

$$\lambda_n(\widehat{\beta}_n) = [\frac{1}{n}\sum_t g_t(\widehat{\beta}_n)g_t^\tau(\widehat{\beta}_n)]^{-1} \cdot [\frac{1}{n}\sum_t g_t(\widehat{\beta}_n)] + b_n(\widehat{\beta}_n)$$

where $\|b_n(\widehat{\beta}_n)\| = o_p(n^{-\frac{1}{2}})$, and $\|\frac{1}{n}\sum_t g_t(\widehat{\beta}_n)\| = O_p(n^{-\frac{1}{2}})$. Then

$$\begin{aligned} & l_n(\widehat{\beta}_n) \\ & = \frac{n}{2}[\frac{1}{n}\sum_t g_t(\widehat{\beta}_n)]^\tau \{[\frac{1}{n}\sum_t g_t(\widehat{\beta}_n)g_t^\tau(\widehat{\beta}_n)]^{-1}[\frac{1}{n}\sum_t g_t(\widehat{\beta}_n)] + b_n(\widehat{\beta}_n)\} + o_p(1) \\ & = \frac{n}{2}[\frac{1}{n}\sum_t g_t(\widehat{\beta}_n)]^\tau \widehat{\lambda}_n + o_p(1). \end{aligned}$$

Thus

$$\begin{aligned} 2l_n(\widehat{\beta}_n) &= n[\frac{1}{n}\sum_t g_t(\widehat{\beta}_n)]^\tau \times \widehat{\lambda}_n + o_p(1) \\ &= n[\frac{1}{n}\sum_t g_t(\beta^{(0)}) + \frac{1}{n}\sum_t \frac{\partial g_t(\beta^{(0)})}{\partial \beta}(\widehat{\beta}_n - \beta^{(0)}) + O_p(n^{-1})]^\tau \widehat{\lambda}_n + o_p(1) \\ &= -nQ_{1n}^\tau(\beta^{(0)}, 0)[\ S_{11}^{-1}(I + S_{12}S_{22.1}^{-1}S_{21}S_{11}^{-1})\ ]Q_{1n}(\beta^{(0)}, 0) + o_p(1). \end{aligned}$$

At the same time, as for $2l_n(\beta^{(0)})$, we have

$$\begin{aligned} 2l_n(\beta^{(0)}) &= n[\frac{1}{n}\sum g_t(\beta^{(0)})]^\tau [\frac{1}{n}\sum g_t(\beta^{(0)})g_t^\tau(\beta^{(0)})]^{-1}[\frac{1}{n}\sum g_t(\beta^{(0)})] + o_p(1) \\ &= -nQ_{1n}^\tau(\beta^{(0)}, 0)S_{11}^{-1}Q_{1n}(\beta^{(0)}, 0) + o_p(1). \end{aligned}$$

Hence

$$\begin{aligned} T =& 2l_n(\beta^{(0)}) - 2l_n(\widehat{\beta}_n) \\ =& nQ_{1n}^\tau(\beta^{(0)}, 0)(S_{11}^{-1} + S_{11}^{-1}S_{12}S_{22.1}^{-1}S_{21}S_{11}^{-1} - S_{11}^{-1})Q_{1n}(\beta^{(0)}, 0) + o_p(1) \\ =& [(-S_{11})^{-\frac{1}{2}}\sqrt{n}Q_{1n}(\beta^{(0)}, 0)]^\tau \times [(-S_{11})^{-\frac{1}{2}}S_{12}S_{22.1}^{-1}S_{21}(-S_{11})^{-\frac{1}{2}}] \times \\ & [(-S_{11})^{-\frac{1}{2}}\sqrt{n}Q_{1n}(\beta^{(0)}, 0)] + o_p(1). \end{aligned}$$

Since

$$R \triangleq (-S_{11})^{-\frac{1}{2}}S_{12}S_{22.1}^{-1}S_{21}(-S_{11})^{-\frac{1}{2}} = R^2$$

and

$$\mathrm{tr}(R) = \mathrm{tr}\{(-S_{11})^{-\frac{1}{2}}S_{12}S_{22.1}^{-1}S_{21}(-S_{11})^{-\frac{1}{2}}\} = q,$$

we have

$$T \longrightarrow \chi_q^2.$$

□

References

[1] BROcKWELL P J, DAVIS R A. *Time Series: Theory and Methods (II)* [J]. Springer-Verlag, New York, 1991.

[2] CHENG Shihong, *Advanced Probability Theorem* [M]. Beijing: Peking Univ Press, 2000.

[3] CHEN Xia. *The law of the iterated logarithm for m-dependent Banach space valued random variables* [J]. J. Theoret. Probab., 1997, **10**(3): 695–732.

[4] FAN Jinqin, YAO Qiwei. *Nonlinear Time Series: Nonparametric and Parametric Methods* [M]. Springer-Verlag, New York, 2002.

[5] HUANG Xiaowei, SONG Lixin. *Empirical likelihood inference for the β-ARCH model* [J]. J. Systems Sci. Math. Sci., 2007, **27**(1): 113–123. (in Chinese)

[6] HUANG Xiaowei. *On the Research of Paramter Estimation for a class of β-ARCH model* [D]. PhD Paper, Changchun, 2004.

[7] KITAMURA Y. *Empirical likelihood methods with weakly dependent processes* [J]. Ann. Statist., 1997, **25**(5): 2084–2102.

[8] LU Chuanrong, LIN Zhengyan. *Limit Theorem for Mixing Dependent Variables* [M]. Beijing: Science Press, 1997.

[9] MONTI A C. *Empirical likelihood confidence regions in time series models* [J]. Biometrika, 1997, **84**(2): 395–405.

[10] OWEN A B. *Empirical likelihood ratio confidence intervals for a single functional* [J]. Biometrika, 1988, **75**(2): 237–249.

[11] OWEN A. *Empirical likelihood ratio confidence regions* [J]. Ann. Statist., 1990, **18**(1): 90–120.

[12] OWEN A. *Empirical likelihood for linear models* [J]. Ann. Statist., 1991, **19**(4): 1725–1747.

[13] QIN Jing, LAWLESS J. *Empirical likelihood and general estimating equations* [J]. Ann. Statist., 1994, **22**(1): 300–325.

[14] ROCKFELLAR R T. *Convex Analysis* [M]. Princeton University Press, Princeton, N. J., 1970.

[15] SERFLING R J. *Approximation Theorems of Mathematical Statistics* [M]. John Wiley & Sons, Inc., New York, 1980.