EMPIRICAL LIKELIHOOD APPROACH FOR NON GAUSSIAN STATIONARY PROCESSES

Hiroaki OGATA

Received August 6, 2005; revised August 26, 2005

Abstract. For a class of non Gaussian stationary processes, we develop the empirical likelihood approach. For this it is known that Whittle likelihood is the most fundamental tool to get a good estimator of unknown parameter, and that the score functions are asymptotically chi-square distributed. Motivated by the Whittle likelihood, we apply the empirical likelihood approach to its derivative. This paper provides a rigorous proof on convergence of our empirical likelihood to a chi-square distribution. Also, some numerical studies on confidence region will be given.

1. Introduction. Empirical likelihood method is used when the distribution of an appropriate pivotal quantity is unknown. It is shown that empirical likelihood ratio is asymptotically chi-square distributed (e.g. Owen (2001)). However, most of studies on this topic are aimed to independent data.

For dependent sample, Monti (1997) applied the empirical likelihood approach to the derivative of the Whittle likelihood, and showed that the empirical likelihood ratio is asymptotically χ^2-distributed. The results were applied to the problem of testing and construction of a confidence region.

Although Monti's results are innovative in time series analysis, the theoretical proofs of the asymptotic results essentially rely on the circular Gaussian assumption for the concerned process like as Anderson (1977). Therefore this paper provides a rigorous proof for asymptotics of the empirical Whittle likelihood ratio using the non Gaussian and dependent structure essentially. Also, some numerical studies on confidence region will be given.

This paper is organized as follows. Section 2 describes our setting and preliminary results for the periodogram. In Section 3, we explain the empirical likelihood approach for the Wittle likelihood. The asymptotic distribution of the empirical likelihood ratio is derived. Section 4 provides some numerical studies on confidence region based on our results. The proof of theorem is relegated to Section 5.

As for notations used in this paper, we denote the convergence in probability by $\xrightarrow{p}$, the convergence in distribution by $\xrightarrow{d}$, the set of all integers by $\mathbf{Z}$, and Kronecker's delta by $\delta(m,n)$.

2. Setting and Preliminaries. We consider a scalar-valued linear process $\{X(t); t \in \mathbf{Z}\}$, generated as

$$X(t) = \sum_{j=0}^{\infty} G(j)\, e(t-j), \qquad t \in \mathbf{Z}, \tag{2.1}$$

where $\{e(t)\}$ is a sequence of random variables satisfying $E\{e(t)\} = 0$ and $E\{e(t)e(s)\} = \delta(t,s)\sigma^2$, with $\sigma^2 > 0$, $G(j)$'s are constants, and the X, e and G are all real. If $\sum_{j=0}^{\infty} G(j)^2 <$

2000 *Mathematics Subject Classification* . 62G30; 62F25; 62M15.
Key words and phrases. Empirical likelihood, spectral density, Whittle likelihood, Non Gaussian stationary process, Estimating function.

∞ (this condition is assumed throughout), the process $\{X(t)\}$ is a second-order stationary process, and has the spectral density function

$$f_{XX}(\omega) = \frac{\sigma^2}{2\pi}\Big|\sum_{j=0}^{\infty} G(j)e^{-i\omega j}\Big|^2, \qquad -\pi \le \omega \le \pi. \tag{2.2}$$

For the stretch $X(t),\ t = 1, \dots, T$, we denote by $I_{XX}(\omega)$, the periodogram; namely

$$I_{XX}(\omega) = \frac{1}{2\pi T}|d_X(\omega)|^2, \qquad \text{where} \qquad d_X(\omega) = \sum_{t=1}^{T} X(t)\exp\{-i\omega t\} \quad -\pi < \omega < \pi.$$

We set down the following assumptions.

Assumption 2.1.

(i) $\{X(t)\}$ *is kth-order stationary with all of whose moments exist.*

(ii) The joint kth-order cumulant $c_{X^k}(u_1, \dots, u_{k-1})$ *of* $X(t), X(t+u_1), \dots, X(t+u_{k-1})$ *satisfies*

$$\sum_{u_1,\dots,u_{k-1}=-\infty}^{\infty} [1 + |u_j|]\,|c_{X^k}(u_1, \dots, u_{k-1})| < \infty$$

for $j = 1, \dots, k-1$ *and any* k, $\ k = 2, 3, \dots$.

Assumption 2.2. *For the sequence* $\{C_k\}$ *defined by*

$$C_k = \sum_{u_1,\dots,u_k=-\infty}^{\infty} |c_{X^k}(u_1, \dots, u_{k-1})|,$$

it holds that

$$\sum_{k=1}^{\infty} C_k z^k/k! < \infty$$

for z *in a neighborhood of 0.*

Denote by $f_{X^k}(\omega_1, \dots, \omega_{k-1})$, the kth-order spectral density of the process $\{X(t)\}$; namely

$$f_{X^k}(\omega_1, \dots, \omega_{k-1}) = (2\pi)^{-k+1} \sum_{u_1,\dots,u_k=-\infty}^{\infty} c_{X^k}(u_1, \dots, u_{k-1}) \exp\Big\{-i\sum_{j=1}^{k-1} u_j\omega_j\Big\}.$$

In what follows, we state the fundamental results on periodogram, which will be used in the next section.

Lemma 2.1. *Let* $\{X(t)\}$ *satisfy Assumption 2.1. Let* $\boldsymbol{A}(\omega), -\pi \le \omega < \pi$ *be a q-dimensional vector valued continuous function, satisfying* $\boldsymbol{A}(\omega) = \boldsymbol{A}(-\omega)$*. Then*

$$T^{-\frac{1}{2}} \sum_{t=1}^{T} \boldsymbol{A}(\lambda_t)\{I_{XX}(\lambda_t) - EI_{XX}(\lambda_t)\} \xrightarrow{d} N\left(0, \Sigma^{(1)}\right) \quad (T \to \infty),$$

where $\lambda_t = 2\pi t/T$ *(throughout this paper) and*

$$\Sigma^{(1)} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \boldsymbol{A}(\alpha)\boldsymbol{A}(\beta)' f_{X^4}(\alpha,-\alpha,\beta)\ d\alpha d\beta + \frac{1}{\pi}\int_{-\pi}^{\pi} \boldsymbol{A}(\alpha)\boldsymbol{A}(\alpha)' f_{XX}(\alpha)^2\ d\alpha.$$

Lemma 2.2. *Under the same assumption as in Lemma 2.1, it holds that*

$$T^{-1}\sum_{t=1}^{T}\{\boldsymbol{A}(\lambda_t)I_{XX}(\lambda_t)\}\{\boldsymbol{A}(\lambda_t)I_{XX}(\lambda_t)\}' \xrightarrow{p} \Sigma^{(2)} \quad (T\to\infty),$$

where

$$\Sigma^{(2)} = \frac{1}{\pi}\int_{-\pi}^{\pi} \boldsymbol{A}(\alpha)\boldsymbol{A}(\alpha)' f_{XX}(\alpha)^2\ d\alpha.$$

The proofs of Lemmas 2.1 and 2.2 are essentially given by Brillinger (1981, Theorem 5.10.1) or Hosoya and Taniguchi (1982).

3. Empirical likelihood approach for time series. Empirical likelihood is a nonparametric method of statistical inference. It allows us to use likelihood methods, without assuming that the data come from a known family of distribution. Empirical likelihood method is based on the nonparametric likelihood ratio $R(F) = \prod_{i=1}^{n} nw_i$ where F is an arbitrary distribution which has probability w_i on the data $\boldsymbol{X}_i$. We use this ratio $R(F)$ as a basis for hypothesis testing and confidence intervals.

When we are interested in parameter $\boldsymbol{\theta}\in\mathbf{R}^q$ which satisfies $E[\boldsymbol{m}(\boldsymbol{X},\boldsymbol{\theta})] = \boldsymbol{0}$, where $\boldsymbol{m}(\boldsymbol{X},\boldsymbol{\theta})\in\mathbf{R}^q$ is the vector-valued function, called estimating function, we consider the empirical likelihood ratio function $\mathcal{R}(\boldsymbol{\theta})$ (defined in (3.3) below). As a test statistic, it is shown that $-2\log\mathcal{R}(\boldsymbol{\theta})$ tends to chi-square with degree of freedom q, when $\boldsymbol{X}_i's$ have identically independent distribution, (e.g. Owen (2001)).

Here, we consider the case of dependent sample. When $\{X(t)\}$ is a Gaussian circular ARMA process, Anderson (1977) showed that the log likelihood for $\boldsymbol{X} = (X(1),\ldots,X(T))'$ becomes, disregarding a constant term,

$$LL_c(\boldsymbol{\theta}) = -\sum_{t=1}^{T}\left\{\log f_{XX}(\lambda_t;\boldsymbol{\theta}) + \frac{I_{XX}(\lambda_t)}{f_{XX}(\lambda_t;\boldsymbol{\theta})}\right\},$$

and that $2I_{XX}(\lambda_t)/f_{XX}(\lambda_t;\boldsymbol{\theta}), t = 1,\ldots,(T/2)-1$ or $(T-1)/2$, are independently distributed, each with a χ_2^2-distribution, where $I_{XX}(\lambda)$ is the periodogram of $\boldsymbol{X}$ and $f_{XX}(\lambda;\boldsymbol{\theta})$ is the spectral density. Without the assumption of circular Gaussian ARMA process, it is known that Anderson's results hold asymptotically (e.g. Taniguchi and Kakizawa (2000)). That is, if $\{X(t)\}$ is an appropriate stationary process, $2I_{XX}(\lambda_t)/f_{XX}(\lambda_t;\boldsymbol{\theta}),\ t = 1,\ldots,(T/2)-1$ or $(T-1)/2$ are asymptotically independent and asymptotically χ_2^2-distributed.

Even if Gaussianity is assumed, without circular assumption, $2I_{XX}(\lambda_t)/f_{XX}(\lambda_t;\boldsymbol{\theta})$'s are not i.i.d. χ_2^2, exactly. Therefore when $\{X(t)\}$ is a non-Gaussian process, it is valueable to consider the empirical likelihood. Monti (1997) applied the spectral approach of this type to the empirical likelihood, and considerd an integral version of $LL_c(\boldsymbol{\theta})$, which is called the Whittele likelihood, that is,

$$W(\boldsymbol{\theta}) \equiv \int_{-\pi}^{\pi} \left\{ \log f_{XX}(\omega;\boldsymbol{\theta}) + \frac{I_{XX}(\omega)}{f_{XX}(\omega;\boldsymbol{\theta})} \right\} d\omega, \tag{3.1}$$

and used $\psi_t(\boldsymbol{\theta}) = (\partial/\partial\boldsymbol{\theta}) \{\log f_{XX}(\lambda_t;\boldsymbol{\theta}) + I_{XX}(\lambda_t)/f_{XX}(\lambda_t;\boldsymbol{\theta})\}$ as a counterpart of Owen's $\boldsymbol{m}(\boldsymbol{X},\boldsymbol{\theta})$. Then, Monti (1997) showed that $-2\log\mathcal{R}(\boldsymbol{\theta})$ tends to chi-square with degree of freedom q. However, the proof of the above result essensially relies on Anderson's results.

In this paper, assuming that $\{X(t)\}$ is a non-Gaussian stationary process, we give the rigorous proof of it. First, we impose the following assumptions.

Assumption 3.1. *$f_{XX}(\omega;\boldsymbol{\theta})$ is continuously twice differentiable with respect to $\boldsymbol{\theta}$.*

Assumption 3.2.

(i) $\boldsymbol{\theta}_0$ is the true parameter of $\boldsymbol{\theta}$.

(ii) $\boldsymbol{\theta}_0$ is innovation free, that is,

$$\int_{-\pi}^{\pi} \frac{\partial}{\partial\boldsymbol{\theta}} \{f_{XX}(\omega;\boldsymbol{\theta})\}^{-1} f_{XX}(\omega;\boldsymbol{\theta})\ d\omega \,\Big|_{\theta=\theta_0} = \mathbf{0}. \tag{3.2}$$

If $\boldsymbol{\theta}$ is innovation-free, $(\partial/\partial\boldsymbol{\theta})W(\boldsymbol{\theta}) = 0$ becomes $\int_{-\pi}^{\pi}(\partial/\partial\boldsymbol{\theta})\{I_{XX}(\omega)/f_{XX}(\omega;\boldsymbol{\theta})\}\, d\omega\,|_{\theta=\theta_0} = \mathbf{0}$ and its discriterized version becomes $\sum_{t=1}^{T}(\partial/\partial\boldsymbol{\theta})\{I_{XX}(\lambda_t)/f_{XX}(\lambda_t;\boldsymbol{\theta})\}\,|_{\theta=\theta_0} = \mathbf{0}$. Because it is known that $E[I_{XX}(\lambda_t)]$ converges to $f_{XX}(\lambda_t;\boldsymbol{\theta})$, we can see that $E[(\partial/\partial\boldsymbol{\theta})\{I_{XX}(\lambda_t)/f_{XX}(\lambda_t;\boldsymbol{\theta})\}\,|_{\theta=\theta_0}] \to 0$, which motivates our empirical likelihood ratio function $\mathcal{R}(\boldsymbol{\theta})$ defined by

$$\mathcal{R}(\boldsymbol{\theta}) = \max\left\{ \prod_{t=1}^{T} T w_t \mid \sum_{t=1}^{T} w_t \boldsymbol{m}(\lambda_t;\boldsymbol{\theta}) = \mathbf{0},\ w_t \geq 0,\ \sum_{t=1}^{T} w_t = 1 \right\}. \tag{3.3}$$

We set down the following further assumption.

Assumption 3.3. *The process $\{e(t)\}$ satisfies*

$$cum\{e(t_1), e(t_2), e(t_3), e(t_4)\} = \begin{cases} \kappa^4 & (t_1 = t_2 = t_3 = t_4) \\ 0 & (otherwise) \end{cases}.$$

Then we get the following theorem. The proof is given in Section 5.

Theorem 3.1. *Let $\{X(t)\}$ be a scalar-valued linear process defined in (2.1), and satisfy Assumptions 2.1 $\sim$ 2.2 and 3.1 $\sim$ 3.3. Then $-2\log\mathcal{R}(\boldsymbol{\theta}_0) \xrightarrow{d} \chi^2_{(q)}$ as $T \to \infty$, where*

$$\boldsymbol{m}(\lambda_t;\boldsymbol{\theta}) = \frac{\partial}{\partial\boldsymbol{\theta}} \left\{ \frac{I_{XX}(\lambda_t)}{f_{XX}(\lambda_t;\boldsymbol{\theta})} \right\}.$$

Using this theorem, we can construct a confidence regions on $\boldsymbol{\theta}$. First, we choose a proper threshold value z_α, which is α percentail of χ^2_2. Then we caluculate $-2\log\mathcal{R}(\boldsymbol{\theta})$ at numerous points over the range and consutruct the region

$$C_{\alpha,T} = \{\boldsymbol{\theta} \mid -2\log\mathcal{R}(\boldsymbol{\theta}) > z_\alpha\}.$$

4 Examples. Let us consider the following ARMA(1,1) model

$$X(t) + \beta X(t-1) = e(t) + \alpha e(t-1),$$

where $e(t)$'s are i.i.d. $N(0, \sigma^2)$. We set $\alpha_0 = 0.2$, $\beta_0 = 0.35$, $\sigma^2 = 1$ and the length of observations $T = 300$. Let $(\hat{\alpha}_{ML}, \hat{\beta}_{ML})$ be the quasi-maximum likelihood estimator of (α_0, β_0) which minimizes $W(\boldsymbol{\theta})$ in (3.1). Then it is known that

$$\sqrt{T}\begin{pmatrix} \hat{\alpha}_{ML} - \alpha_0 \\ \hat{\beta}_{ML} - \beta_0 \end{pmatrix} \xrightarrow{d} N\left(\mathbf{0}, \begin{pmatrix} \frac{1}{1-\alpha_0{}^2} & -\frac{1}{1-\alpha_0\beta_0} \\ -\frac{1}{1-\alpha_0\beta_0} & \frac{1}{1-\beta_0{}^2} \end{pmatrix}^{-1}\right), \tag{4.1}$$

(e.g. Hosoya and Taniguchi (1982)). Then we can construct a confidence region by (4.1). Figure 1 shows the 90% empirical likelihood confidence rigion (solid line) and the usual confidence region based on the maximum likelihood estimator (dashed line) for (α, β). Empirical confidence gives the larger region than usual confidence. However, the usual method must specify the parametric family to use though we might not know it. Such misspecification can cause the confidence region to be failed completely. Therehore, the empirical likelihood method is worth consideration.

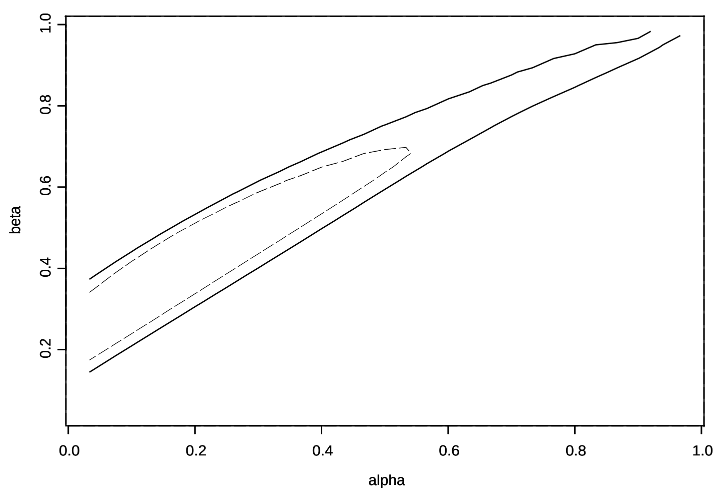

Figure 1:

5. Proofs. In this section we give the proof of Theorem 3.1. For this we need the following lemma which is due to Brillinger (2001, Theorem 4.5.1).

Lamma A5.1. *Let $\{X(t),\ t \in \mathbf{Z}\}$ satisfy Assumption 2.2 and have mean 0. Then*

$$\overline{\lim_{T\to\infty}} \sup_{\omega} \frac{|d_X(\omega)|}{(T \log T)^{1/2}} \le 2\{2\pi \sup_{\omega} f_{XX}(\omega)\}^{1/2}$$

with probability 1.

Using the above Lemma, we prove Theorem 3.1.

***Proof of Theorem 3.1*.**
To find the maximizing weights $w_t's$ of (3.3), we proceed by the Lagrange multiplier method. Write

$$G = \sum_{t=1}^{T} \log(Tw_t) - T\boldsymbol{\phi}' \sum_{t=1}^{T} w_t \boldsymbol{m}(\lambda_t; \boldsymbol{\theta}_0) + \gamma \left(\sum_{t=1}^{T} w_t - 1 \right),$$

where $\boldsymbol{\phi} \in \mathbf{R}^q$ and $\gamma \in \mathbf{R}$ are Lagrange multipliers. Setting $\partial G / \partial w_t = 0$ gives

$$\frac{\partial G}{\partial w_t} = \frac{1}{w_t} - T\boldsymbol{\phi}' \boldsymbol{m}(\lambda_t; \boldsymbol{\theta}_0) + \gamma = 0.$$

So, the equation $\sum_{t=1}^{T} w_t(\partial G / \partial w_t) = 0$ gives $\gamma = -T$. Then, we may write

$$w_t = \frac{1}{T} \frac{1}{1 + \boldsymbol{\phi}' \ \boldsymbol{m}(\lambda_t; \boldsymbol{\theta}_0)}.$$

where the vector $\boldsymbol{\phi} = \boldsymbol{\phi}(\boldsymbol{\theta}_0)$ satisfies q equations given by

$$\boldsymbol{g}(\boldsymbol{\phi}) := \frac{1}{T} \sum_{t=1}^{T} \frac{\boldsymbol{m}(\lambda_t; \boldsymbol{\theta}_0)}{1 + \boldsymbol{\phi}' \ \boldsymbol{m}(\lambda_t; \boldsymbol{\theta}_0)} = \mathbf{0}. \tag{5.1}$$

Let $\boldsymbol{\phi} = \|\boldsymbol{\phi}\| \boldsymbol{u}$ where $\boldsymbol{u} \in U$, a set of unit vector. Introduce

$$Y_t := \boldsymbol{\phi}' \boldsymbol{m}(\lambda_t; \boldsymbol{\theta}_0), \quad \textit{and} \quad Z_T^* := \max_{1 \le t \le T} \|\boldsymbol{m}(\lambda_t; \boldsymbol{\theta}_0)\|.$$

Substituting $1/(1 + Y_t) = 1 - Y_t/(1 + Y_t)$ into $\boldsymbol{u}'\boldsymbol{g}(\boldsymbol{\phi}) = 0$ and simplifying, we find that

$$\boldsymbol{u}' \left\{ \frac{1}{T} \sum_{t=1}^{T} \left(1 - \frac{Y_t}{1 + Y_t} \right) \boldsymbol{m}(\lambda_t; \boldsymbol{\theta}_0) \right\} = 0$$

$$\boldsymbol{u}' \left(\frac{1}{T} \sum_{t=1}^{T} \frac{\boldsymbol{\phi}' \boldsymbol{m}(\lambda_t; \boldsymbol{\theta}_0)}{1 + Y_t} \boldsymbol{m}(\lambda_t; \boldsymbol{\theta}_0) \right) = \boldsymbol{u}' \left(\frac{1}{T} \sum_{t=1}^{T} \boldsymbol{m}(\lambda_t; \boldsymbol{\theta}_0) \right)$$

$$\|\boldsymbol{\phi}\| \boldsymbol{u}' \left(\frac{1}{T} \sum_{t=1}^{T} \frac{\boldsymbol{m}(\lambda_t; \boldsymbol{\theta}_0) \boldsymbol{m}(\lambda_t; \boldsymbol{\theta}_0)'}{1 + Y_t} \right) \boldsymbol{u} = \boldsymbol{u}' \left(\frac{1}{T} \sum_{t=1}^{T} \boldsymbol{m}(\lambda_t; \boldsymbol{\theta}_0) \right) \tag{5.2}$$

Let

$$S := \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{m}(\lambda_t; \boldsymbol{\theta}_0) \, \boldsymbol{m}(\lambda_t; \boldsymbol{\theta}_0)'.$$

Every $w_t > 0$, so $1 + Y_t > 0$ and therefore by (5.2), we get

$$\begin{aligned}
\|\boldsymbol{\phi}\| \boldsymbol{u}' S \boldsymbol{u} &\le \|\boldsymbol{\phi}\| \boldsymbol{u}' \left(\frac{1}{T} \sum_{t=1}^{T} \frac{\boldsymbol{m}(\lambda_t; \boldsymbol{\theta}_0) \boldsymbol{m}(\lambda_t; \boldsymbol{\theta}_0)'}{1 + Y_t} \right) \boldsymbol{u} \cdot (1 + \max_t Y_t) \\
&\le \|\boldsymbol{\phi}\| \boldsymbol{u}' \left(\frac{1}{T} \sum_{t=1}^{T} \frac{\boldsymbol{m}(\lambda_t; \boldsymbol{\theta}_0) \boldsymbol{m}(\lambda_t; \boldsymbol{\theta}_0)'}{1 + Y_t} \right) \boldsymbol{u} \cdot (1 + \|\boldsymbol{\phi}\| Z_t^*) \\
&= \boldsymbol{u}' \left(\frac{1}{T} \sum_{t=1}^{T} \boldsymbol{m}(\lambda_t; \boldsymbol{\theta}_0) \right) (1 + \|\boldsymbol{\phi}\| Z_t^*).
\end{aligned} \tag{5.3}$$

Then by (5.3), we get

$$\|\boldsymbol{\phi}\| \left[\boldsymbol{u}' S \boldsymbol{u} - Z_T^* \boldsymbol{u}' \left\{\frac{1}{T}\sum_{t=1}^{T} \boldsymbol{m}(\lambda_t;\boldsymbol{\theta}_0)\right\}\right] \leq \boldsymbol{u}' \left\{\frac{1}{T}\sum_{t=1}^{T} \boldsymbol{m}(\lambda_t;\boldsymbol{\theta}_0)\right\}. \tag{5.4}$$

Now, we evaluate the order of $(1/T)\sum_{t=1}^{T} \boldsymbol{m}(\lambda_t;\boldsymbol{\theta}_0)$, S and Z_T^*. When $\boldsymbol{\theta}_0$ is innovation-free,

$$\begin{aligned}
& \frac{1}{\sqrt{T}}\sum_{t=1}^{T} \frac{\partial}{\partial\boldsymbol{\theta}}\{f_{XX}(\lambda_t;\boldsymbol{\theta})\}^{-1}\Big|_{\theta=\theta_0} EI_{XX}(\lambda_t) \\
= \; & \frac{\sqrt{T}}{2\pi}\,\frac{2\pi}{T}\sum_{t=1}^{T} \frac{\partial}{\partial\boldsymbol{\theta}}\{f_{XX}(\lambda_t;\boldsymbol{\theta})\}^{-1}\Big|_{\theta=\theta_0} EI_{XX}(\lambda_t) \\
= \; & \frac{\sqrt{T}}{2\pi}\left\{\int_{-\pi}^{\pi} \frac{\partial}{\partial\boldsymbol{\theta}}\{f_{XX}(\omega;\boldsymbol{\theta})\}^{-1}\Big|_{\theta=\theta_0} f_{XX}(\omega;\boldsymbol{\theta})\; d\omega + O(T^{-1})\right\} \\
= \; & O(T^{-\frac{1}{2}}). \qquad (5.5)
\end{aligned}$$

In Lemma 2.1, let $\boldsymbol{A}(\omega) = (\partial/\partial\boldsymbol{\theta}) f_{XX}(\omega;\boldsymbol{\theta})^{-1}|_{\theta=\theta_0}$. Then by (5.5)

$$T^{-\frac{1}{2}}\sum_{t=1}^{T} \frac{\partial}{\partial\boldsymbol{\theta}}\{f_{XX}(\omega;\boldsymbol{\theta})\}^{-1}\Big|_{\theta=\theta_0} I_{XX}(\lambda_t) \xrightarrow{d} N(0,\Sigma^{(1)}),$$

which implies

$$\frac{1}{T}\sum_{t=1}^{T} \boldsymbol{m}(\lambda_t;\boldsymbol{\theta}_0) = O_p(T^{-\frac{1}{2}}). \tag{5.6}$$

Similarly, letting $(\partial/\partial\boldsymbol{\theta}) f_{XX}(\omega;\boldsymbol{\theta})^{-1}|_{\theta=\theta_0} = \boldsymbol{A}(\omega)$ in Lemma 2.2, we get

$$T^{-1}\sum_{t=1}^{T}\left\{\frac{\partial}{\partial\boldsymbol{\theta}}\{f_{XX}(\omega;\boldsymbol{\theta})\}^{-1}\Big|_{\theta=\theta_0} I_{XX}(\lambda_t)\right\}\left\{\frac{\partial}{\partial\boldsymbol{\theta}}\{f_{XX}(\omega;\boldsymbol{\theta})\}^{-1}\Big|_{\theta=\theta_0} I_{XX}(\lambda_t)\right\}' \xrightarrow{p} \Sigma^{(2)},$$

which implies

$$\frac{1}{T}\sum_{t=1}^{T} \boldsymbol{m}(\lambda_t;\boldsymbol{\theta}_0)\boldsymbol{m}(\lambda_t;\boldsymbol{\theta}_0)' = S = O_p(1). \tag{5.7}$$

It follows from Lemma A5.1 that

$$\begin{aligned}
Z_T^* \; &= \; \max_{1\leq t\leq T} \|\boldsymbol{A}(\lambda_t)\|\;|I_{XX}(\lambda_t)| \\
&\leq \; \sup_{\alpha}\|\boldsymbol{A}(\alpha)\|\cdot(2\pi T)^{-1}\left(\sup_{\alpha}|d_X(\alpha)|\right)^2 \\
&= \; O(\log T). \qquad (5.8)
\end{aligned}$$

From (5.4)-(5.8), it is seen that

$$\|\boldsymbol{\phi}\|\left[O_p(1) - O(\log T)O_p(T^{-1/2})\right] \leq O_p(T^{-1/2}).$$

Therefore,

$$\|\boldsymbol{\phi}\| = O_p(T^{-1/2}).$$

While, we have from (5.8) that

$$\max_{1\le t\le T} |Y_t| = O_p(T^{-\frac{1}{2}})O(\log T) = O_p(T^{-\frac{1}{2}}\log T), \tag{5.9}$$

and from (5.1) that

$$\begin{aligned}\mathbf{0} &= \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{m}(\lambda_t;\boldsymbol{\theta}_0)\frac{1}{1+Y_t} \\ &= \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{m}(\lambda_t;\boldsymbol{\theta}_0)\left(1 - Y_t + \frac{Y_t^2}{1+Y_t}\right) \\ &= \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{m}(\lambda_t;\boldsymbol{\theta}_0) - S\boldsymbol{\phi} + \frac{1}{T}\sum_{t=1}^{T}\frac{\boldsymbol{m}(\lambda_t;\boldsymbol{\theta}_0)Y_t^2}{1+Y_t}. \qquad (5.10)\end{aligned}$$

Noting that

$$\frac{1}{T}\sum_{t=1}^{T}\|\boldsymbol{m}(\lambda_t;\boldsymbol{\theta}_0)\|^3 \le \frac{1}{T}\sum_{t=1}^{T} Z_T^*\|\boldsymbol{m}(\lambda_t;\boldsymbol{\theta}_0)\|^2 = O(\log T),$$

we can see that the final term in (5.10) has a norm bounded by

$$\frac{1}{T}\sum_{t=1}^{T}\|\boldsymbol{m}(\lambda_t;\boldsymbol{\theta}_0)\|^3\|\boldsymbol{\phi}\|^2|1+Y_t|^{-1} = O(\log T)\ O_p(T^{-1})\ O_p(1) = O_p(T^{-1}\log T),$$

hence, we can write

$$\boldsymbol{\phi} = S^{-1}\left\{\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{m}(\lambda_t;\boldsymbol{\theta}_0)\right\} + \boldsymbol{\beta},$$

where $\boldsymbol{\beta} = O_p(T^{-1}\log T)$.

By (5.9), we may write

$$\log(1+Y_t) = Y_t - \frac{1}{2}Y_t^2 + \eta_t$$

where for some finite $B > 0$

$$\Pr(|\eta_t| \le B|Y_t|^3, \quad 1 \le t \le T) \to 1 \quad (T\to\infty).$$

We may write

$$\begin{aligned}-2\log\mathcal{R}(\boldsymbol{\theta}_0) &= -2\sum_{t=1}^{T}\log(Tw_t) = 2\sum_{t=1}^{T}\log(1+Y_t) \\ &= 2\sum_{t=1}^{T}Y_t - \sum_{t=1}^{T}Y_t^2 + 2\sum_{t=1}^{T}\eta_t \\ &= \left\{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\boldsymbol{m}(\lambda_t;\boldsymbol{\theta}_0)\right\}' S^{-1}\left\{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\boldsymbol{m}(\lambda_t;\boldsymbol{\theta}_0)\right\} - T\boldsymbol{\beta}'S\boldsymbol{\beta} + 2\sum_{t=1}^{T}\eta_t \\ &= (A) - (B) + (C). \qquad \text{(say)}\end{aligned}$$

Here it is seen

$$\begin{aligned}(B) &= TO_p(T^{-1}\log T)O_p(1)O_p(T^{-1}\log T) = O_p(T^{-1}(\log T)^2),\\ (C) &\leq B\|\boldsymbol{\phi}\|^3\sum_{t=1}^{T}\|\boldsymbol{m}(\lambda_t;\boldsymbol{\theta}_0)\|^3 = O_p(T^{-\frac{3}{2}})O(T\log T) = O_p(T^{-\frac{1}{2}}\log T).\end{aligned}$$

Next, we show that $\Sigma^{(1)} = \Sigma^{(2)}$ when $\boldsymbol{A}(\omega) = (\partial/\partial\boldsymbol{\theta})f_{XX}(\omega;\boldsymbol{\theta})^{-1}|_{\theta=\theta_0}$ and $\boldsymbol{\theta}_0$ is innovation-free. By Assumption 3.3, it holds that

$$\begin{aligned}&f_{XXXX}(\alpha,-\alpha,\beta;\boldsymbol{\theta}_0)\\ =\ &(2\pi)^{-3}\kappa^4\left(\sum_{j=0}^{\infty}G(j;\boldsymbol{\theta}_0)e^{-i\alpha j}\right)\left(\sum_{j=0}^{\infty}G(j;\boldsymbol{\theta}_0)e^{-i\beta j}\right)\left(\sum_{j=0}^{\infty}G(j;\boldsymbol{\theta}_0)e^{i\alpha j}\right)\left(\sum_{j=0}^{\infty}G(j;\boldsymbol{\theta}_0)e^{i\beta j}\right)\\ =\ &(2\pi)^{-1}(\sigma^2)^{-2}\kappa^4 f_{XX}(\alpha;\boldsymbol{\theta}_0)f_{XX}(\beta;\boldsymbol{\theta}_0).\end{aligned}$$

Then

$$\begin{aligned}&\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\boldsymbol{A}(\alpha)\boldsymbol{A}(\beta)' f_{XXXX}(\alpha,-\alpha,\beta;\boldsymbol{\theta}_0)\ d\alpha d\beta\\ =\ &(2\pi)^{-1}(\sigma^2)^{-2}\kappa^4\\ &\times\left[\int_{-\pi}^{\pi}\left\{\frac{\partial}{\partial\boldsymbol{\theta}}f_{XX}(\alpha;\boldsymbol{\theta})^{-1}\Big|_{\theta=\theta_0}\right\}f_{XX}(\alpha;\boldsymbol{\theta}_0)\ d\alpha\right]\\ &\times\left[\int_{-\pi}^{\pi}\left\{\frac{\partial}{\partial\boldsymbol{\theta}}f_{XX}(\beta;\boldsymbol{\theta})^{-1}\Big|_{\theta=\theta_0}\right\}f_{XX}(\beta;\boldsymbol{\theta}_0)\ d\beta\right]'\\ =\ &\mathbf{0}. \qquad (5.11)\end{aligned}$$

From Lemmas 2.1 $\sim$ 2.2 and (5.11), we can see that $\Sigma^{(1)} = \Sigma^{(2)}$. Finally by the central limit theorem of Lemma 2.1, we can show that $(A) \xrightarrow{d} \chi^2_{(q)}$. □

Acknowledgements The author is grateful to Professor M. Taniguchi for his instructive advice. Thanks are extended to a referee for his/her helpful comments.

References

[1] Anderson, T. W. (1977). Estimation for autoregressive moving average models in the time and frequency domains. *Ann. Statist.* **5**, 842-865.

[2] Brillinger,D.R. (1981). *Time Series: Data Analysis and Theory, expanded ed.* Holden-Day, San Francisco.

[3] Hosoya, Y. and Taniguchi, M. (1982). A central limit theorem for stationary processes and the parameter estimation of linear processes. *Ann. Statist.* **10**, 132-153. Correction: (1993) **21** 1115-1117.

[4] Monti, A. C. (1997). Empirical likelihood confidence regions in time series models. *Biometrika.* **84**, 395-405.

[5] Owen, A. B. (2001). *Empirical Likelihood.* Chapman-Hall/CRC, Boca Raton, London, New York, Washinton,D.C.

[6] Taniguchi, M. and Kakizawa, Y. (2000). *Asymptotic Theory of Statistical Inference for Time Series.* Springer-Verlag, New York.

Hiroaki Ogata
Department of Mathematical Sciences
School of Science and Engineering
Waseda University
3-4-1 Okubo Shinjuku-ku
Tokyo 169-8555
Japan
E-mail: hiroakiogata@suou.waseda.jp