POWER PROPERTIES OF EMPIRICAL LIKELIHOOD FOR STATIONARY PROCESSES

Kenichiro TAMAKI*

Received June 22, 2007; revised June 29, 2007

Abstract. In this paper we discuss the second order power properties of empirical likelihood for stationary processes. The asymptotic distribution of empirical likelihood ratio statistics under a sequence of local alternatives is given.

1. Introduction. Since its introduction by Owen (1988, 1990), empirical likelihood has become a useful tool for nonparametric inference. It is well known that the empirical likelihood ratio statistic inherits a number of properties of the parametric likelihood ratio statistic. Owen has shown that the empirical likelihood ratio statistic has limiting chi-squared distribution. Qin and Lawless (1994) connected the theories of empirical likelihood and general estimating equations. Hence using empirical likelihood ratio statistics it is possible to obtain tests and confidence regions for a wide range of problems, including linear models (Owen, 1991). Another property of empirical likelihood which also resembles that of a parametric likelihood is Bartlett correction; see for example Hall and La Scala (1990) for the case of the mean parameters, DiCiccio, et al (1991) for the case of smooth functions of means, Chen (1993, 1994) for linear regression model, Chen and Cui (2006) in the presence of nuisance parameters.

For dependent data, Monti (1997) applied the empirical likelihood approach to the derivative of the Whittle likelihood. Kitamura (1997) considered blockwise empirical likelihood ratios based on data blocks rather than individual observations. Recently, Nordman and Lahiri (2006) introduced a version of empirical likelihood based on the periodogram and spectral estimating equations. They elucidated the asymptotic properties of frequency domain empirical likelihood for linear processes exhibiting both short- and long-range dependence.

In this paper we consider the second order power properties of empirical likelihood for stationary processes. Section 2 provides a survey of frequency domain empirical likelihood which is due to Nordman and Lahiri (2006). Section 3 gives the asymptotic distribution of empirical likelihood ratio statistics under a sequence of local alternatives. The proofs of results are relegated to Section 4.

2. Frequency domain empirical likelihood. Consider inference on a parameter $\theta = (\theta^1, \dots \theta^p)' \in \Theta \subset \mathbf{R}^p$ based on a time stretch $X_1, \dots, X_n$ with spectral density f. We suppose that information about θ exists through a system of general estimating equations. Let

$$G_\theta(\lambda) = (g_{1,\theta}(\lambda), \dots, g_{p,\theta}(\lambda))'.$$

2000 *Mathematics Subject Classification.* 62G30; 62F05; 62M10.
Key words and phrases. empirical likelihood; local power; stationary process.

We assume that G_θ satisfies the spectral moment condition

$$\int_{-\pi}^{\pi} G_{\theta_0}(\lambda) f(\lambda) d\lambda = 0, \tag{1}$$

where θ_0 is the true parameter of θ.

Example 1. Consider interest in the autocorrelation function $\rho(m) = \gamma(m)/\gamma(0)$ at lag $m > 0$, that is $\theta = \rho(m)$. One can select $G_\theta(\lambda) = (e^{im\lambda} + e^{-im\lambda})/2 - \theta$ for autocorrelation inference, because of

$$\int_{-\pi}^{\pi} G_\theta(\lambda) f(\lambda) d\lambda = \gamma(m) - \theta\gamma(0) = 0.$$

To obtain confidence regions for θ, we define

$$R_n(\theta) = \max\left\{ \prod_{j=1}^n n w_j \;\Big|\; \sum_{j=1}^n w_j G_\theta(\lambda_j) I_n(\lambda_j) = 0, w_i \ge 0, \sum_{j=1}^n w_j = 1 \right\},$$

where $\lambda_j = 2\pi j/n$ and $I_n(\lambda) = (2\pi n)^{-1} |\sum_{t=1}^n X_t \exp(it\lambda)|^2$. For given θ, a unique maximum exists, if 0 is inside the convex hull of the point $G_\theta(\lambda_1)I_n(\lambda_1), \ldots, G_\theta(\lambda_n)I_n(\lambda_n)$. An explicit expression for $R_n(\theta)$ can be derived by a Lagrange multiplier argument. Let

$$\mathcal{L} = \sum_{j=1}^n \log(n w_j) - n t' \sum_{j=1}^n w_j G_\theta(\lambda_j) I_n(\lambda_j) + \gamma\left(\sum_{j=1}^n w_j - 1 \right),$$

where $t = (t^1, \ldots, t^p)'$ and γ are Lagrange multipliers. Setting to zero the partial derivative of $\mathcal{L}$ with respect to w_j gives

$$\frac{\partial \mathcal{L}}{\partial w_j} = \frac{1}{w_j} - n t' G_\theta(\lambda_j) I_n(\lambda_j) + \gamma = 0, \quad j = 1, \ldots, n.$$

So

$$0 = \sum_{i=1}^n w_j \frac{\partial \mathcal{L}}{\partial w_j} = n + \gamma,$$

from which $\gamma = -n$. We may therefore write

$$w_j = \frac{1}{n} \frac{1}{1 + t' G_\theta(\lambda_j) I_n(\lambda_j)}, \quad j = 1, \ldots, n.$$

The restriction $\sum_{j=1}^n w_j G_\theta(\lambda_j) I_n(\lambda_j) = 0$ yields p equations

$$\frac{1}{n} \sum_{j=1}^n \frac{G_\theta(\lambda_j) I_n(\lambda_j)}{1 + t' G_\theta(\lambda_j) I_n(\lambda_j)} = 0, \tag{2}$$

from which t can be determined in terms of θ. Thus the empirical log likelihood ratio for θ is defined as

$$\log R_n(\theta) = -\sum_{j=1}^n \log\{1 + t' G_\theta(\lambda_j) I_n(\lambda_j)\}. \tag{3}$$

Since analytic solution of equations (2) and (3) can rarely attained, we have to derive an asymptotic expansion for $-2 \log R_n(\theta)$. Henceforth we assume the following condition.

Assumption 1. (i) $\{X_t\}$ is a real-valued linear process generated by

$$X_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j},$$

where ε_t are independent identically distributed random variables with $E[\varepsilon_t] = 0$, $E[\varepsilon_t^2] = \sigma^2$ and finite absolute moments.

(ii) a_j $(j = 0, \pm 1, \dots,)$ satisfy

$$\sum_{j=0}^{\infty}(1 + |j|^{1+\delta})|a_j| < \infty,$$

for some $\delta > 0$.

(iii) The functions $g_{j,\theta}(\lambda)$ $(j = 1, \dots, p)$ are even and bounded, that is

$$g_{j,\theta}(\lambda) = g_{j,\theta}(-\lambda), \qquad |g_{j,\theta}(\lambda)| < M,$$

for $M > 0$.

(iv) $g_{j,\theta}(\lambda)$ is continuously three times differentiable with respect to θ for $\lambda \in [-\pi, \pi]$.

(v) The $(p \times p)$ matrix $W(\theta_0) = \{W_{ij}(\theta_0)\}$ is positive definite, where

$$W_{ij}(\theta) = \frac{2}{2\pi} \int_{-\pi}^{\pi} g_{i,\theta}(\lambda) g_{j,\theta}(\lambda) f(\lambda)^2 d\lambda.$$

Assumption 1 (ii) ensures that $\{X_t\}$ has the spectral density

$$f(\lambda) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma(j) e^{-ij\lambda},$$

where $\gamma(j) = E[X_t X_{t+j}]$ satisfies

$$\sum_{j=-\infty}^{\infty} (1 + |j|^{1+\delta})|\gamma(j)| < \infty.$$

To simplify the expansions that follow, we define

$$V_{j_1 \cdots j_k}(\theta) = \frac{1}{n} \sum_{l=1}^{n} g_{j_1,\theta}(\lambda_l) \cdots g_{j_k,\theta}(\lambda_l) E[I_n(\lambda_l)^k],$$

$$Z_{j_1 \cdots j_k}(\theta) = \frac{1}{\sqrt{n}} \sum_{l=1}^{n} \{g_{j_1,\theta}(\lambda_l) \cdots g_{j_k,\theta}(\lambda_l) I_n(\lambda_l)^k - V_{j_1 \cdots j_k}(\theta)\}.$$

Henceforth we use the simpler notations $V_{j_1 \cdots j_k}$, $Z_{j_1 \cdots j_k}$, etc. if $V_{j_1 \cdots j_k}(\theta)$, $Z_{j_1 \cdots j_k}(\theta)$, etc. are evaluated at $\theta = \theta_0$. Similarly any function evaluated at the point $\theta = \theta_1$ will be distinguished by the addition of a tilde.

Note that

$$E\{I_n(\lambda_l)\} = f(\lambda_l) - \frac{1}{n} b(\lambda_l) + o(n^{-1}), \tag{4}$$

where

$$b(\lambda) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} |j| \gamma(j) e^{-ij\lambda}.$$

If $\theta = \theta_0 + O(n^{-1/2})$, then, from (1)

$$C_i(\theta) := \frac{1}{\sqrt{n}} \sum_{l=1}^{n} g_{i,\theta}(\lambda_l) E\{I_n(\lambda_l)\} = O(1).$$

Solving (2) gives, using the summation convention,

$$\begin{aligned} t^i = &\frac{1}{\sqrt{n}} V^{ij}(Z_j + C_j) - \frac{1}{n} V^{ij} V^{kl} Z_{jk}(Z_l + C_l) \\ &+ \frac{1}{n} V^{ia} V^{jb} V^{kc} V_{abc}(Z_j + C_j)(Z_k + C_k) + o_p(n^{-1}), \end{aligned} \tag{5}$$

where V^{ij} is the (i,j) component of the inverse matrix of $\{V_{ij}\}$. Substitution of (5) into (3) gives

$$\begin{aligned} -2 \log R_n(\theta) = &V^{ij}(Z_i + C_i)(Z_j + C_j) - \frac{1}{\sqrt{n}} V^{ik} V^{jl} Z_{kl}(Z_i + C_i)(Z_j + C_j) \\ &+ \frac{2}{3\sqrt{n}} V^{ia} V^{jb} V^{kc} V_{abc}(Z_i + C_i)(Z_j + C_j)(Z_k + C_k) + o_p(n^{-1/2}). \end{aligned} \tag{6}$$

3. Local power.

To elucidate the local power of empirical likelihood, we consider testing the null hypothesis that $H : \theta = \theta_0$ against contiguous alternatives $A : \theta = \theta_1 = \theta_0 + n^{-1/2} h$, where $h = (h^1, \ldots, h^p)'$.

First, we give the stochastic expansion of $-2 \log R_n(\theta_0)$ under $\theta = \theta_1$.

Lemma 1. The stochastic expansion of $-2 \log R(\theta_0)$ under $\theta = \theta_1$ is given by

$$\begin{aligned} -2 \log R_n(\theta_0) = &\tilde{W}^{ij}(\tilde{Z}_i - \tilde{A}_{i,a} h^a)(\tilde{Z}_j - \tilde{A}_{j,b} h^b) \\ &+ \frac{2}{\sqrt{n}} \tilde{W}^{ik} \tilde{W}^{jl} \tilde{W}_{kl,c} h^c (\tilde{Z}_i - \tilde{A}_{i,a} h^a)(\tilde{Z}_j - \tilde{A}_{j,b} h^b) \\ &+ \frac{1}{\sqrt{n}} \tilde{W}^{ij}(-2\tilde{Z}_{j,b} h^b + 2\tilde{A}_j + \tilde{A}_{j,bc} h^b h^c)(\tilde{Z}_i - \tilde{A}_{i,a} h^a) \\ &- \frac{1}{\sqrt{n}} \tilde{W}^{ik} \tilde{W}^{jl} \tilde{Z}_{kl}(\tilde{Z}_i - \tilde{A}_{i,a} h^a)(\tilde{Z}_j - \tilde{A}_{j,b} h^b) \\ &+ \frac{2}{3\sqrt{n}} \tilde{W}^{ia} \tilde{W}^{jb} \tilde{W}^{kc} \tilde{W}_{abc}(\tilde{Z}_i - \tilde{A}_{i,a} h^a)(\tilde{Z}_j - \tilde{A}_{j,b} h^b)(\tilde{Z}_k - \tilde{A}_{k,c} h^c) \\ &+ o_p(n^{-1/2}), \end{aligned}$$

where W^{ij} is the (i,j) component of the inverse matrix of $\{W_{ij}\}$,

$$W_{ij,a}(\theta) = \frac{2}{2\pi}\int_{-\pi}^{\pi} g_{i,\theta}(\lambda)\partial_a g_{j,\theta}(\lambda) f(\lambda)^2 d\lambda,$$
$$W_{ijk}(\theta) = \frac{6}{2\pi}\int_{-\pi}^{\pi} g_{i,\theta}(\lambda) g_{j,\theta}(\lambda) g_{k,\theta}(\lambda) f(\lambda)^3 d\lambda,$$
$$Z_{i,a}(\theta) = \frac{1}{\sqrt{n}}\sum_{l=1}^{n} \partial_a g_{i,\theta}(\lambda_l)\{I_n(\lambda_l) - E[I_n(\lambda_l)]\},$$
$$A_i(\theta) = \sum_{l=1}^{n} g_{i,\theta}(\lambda_l) f(\lambda_l) - \frac{1}{2\pi}\int_{-\pi}^{\pi} g_{i,\theta}(\lambda) b(\lambda) d\lambda,$$
$$A_{i,a_1\cdots a_b}(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \partial^b_{a_1\cdots a_b} g_{i,\theta}(\lambda) f(\lambda) d\lambda,$$

and $\partial^b_{a_1\cdots a_b} = \partial^b/\partial\theta^{a_1}\cdots\partial\theta^{a_b}$.

To describe the second order asymptotic properties of $-2\log R_n(\theta_0)$ we use the next lemma.

Lemma 2. (i) $\mathrm{cum}[\tilde{Z}_i, \tilde{Z}_{j,a}] = \tilde{W}_{ij,a} + O(n^{-1})$,

(ii) $\mathrm{cum}[\tilde{Z}_i, \tilde{Z}_{jk}] = \frac{4}{3}\tilde{W}_{ijk} + O(n^{-1})$.

Let $\chi_{p,\Delta}(z)$ is the distribution function for a non-central chi-square variate with degree of freedom p and non-centrality parameter Δ. The following theorem gives the second-order asymptotic expansion of the distribution function of $-2\log R_n(\theta_0)$ under a sequence of local alternatives $\theta = \theta_0 + n^{-1/2}h$.

Theorem 1. The distribution function of $-2\log R_n(\theta_0)$ under a sequence of local alternatives $\theta = \theta_0 + n^{-1/2}h$ has the asymptotic expansion

$$P_{\theta_0+n^{-1/2}h}[-2\log R_n(\theta_0) < z] = \chi_{p,\Delta}(z) + n^{-1/2}\sum_{j=0}^{3} m_j \chi_{p+2j,\Delta}(z) + o(n^{-1/2}),$$

where

$$m_3 = \left(\frac{4}{9}U_{ijk} - \frac{1}{6}M_{ijk}\right) B_{i,a}B_{j,b}B_{k,c}h^a h^b h^c,$$
$$m_2 = \left(-\frac{2}{3}U_{ijk} + \frac{1}{2}M_{ijk}\right) B_{i,a}B_{j,b}B_{k,c}h^a h^b h^c + \left(\frac{4}{3}U_{ijj} - \frac{3}{2}M_{ijj}\right) B_{i,a}h^a,$$
$$m_1 = \left(-\frac{2}{3}U_{ijk}B_{i,a} - \frac{1}{2}M_{ijk}B_{i,a} + U_{jk,a}\right) B_{j,b}B_{k,c}h^a h^b h^c + (M_{ijj} - 2B_i)B_{i,a}h^a,$$
$$m_0 = \left(\frac{2}{9}U_{ijk}B_{i,a}B_{j,b} + \frac{1}{6}M_{ijk}B_{i,a}B_{j,b} + U_{jk,a}B_{j,b} - B_{k,ab}\right) B_{k,c}h^a h^b h^c - \left(\frac{2}{3}U_{ijj} + \frac{1}{2}M_{ijj}\right) B_{i,a}h^a,$$

$\Delta = B_{i,a}B_{i,b}h^a h^b$, and $B_{i,a}$, U_{ijk}, M_{ijk}, etc. are defined in (12) and (14).

Example 2. To observe power properties, we consider the following AR(1) model:

$$X_t - aX_{t-1} = \varepsilon_t, \quad |a| < 1, \quad \varepsilon_t \sim \text{i.i.d. } (0, \sigma^2), \quad E[{\varepsilon_t}^3] = 0.$$

Similarly as Example 1, we consider interest in the autocorrelation function $\rho(m) = a^m$ at lag $m > 0$, that is $\theta = \rho(m)$. One can select $G_\theta(\lambda) = (e^{im\lambda}+e^{-im\lambda})/2-\theta$ for autocorrelation inference. Note that

$$\begin{aligned}\sum_{j=-\infty}^{\infty} \gamma(j)\gamma(m-j) &= \left(\frac{\sigma^2}{1-a^2}\right)^2 a^m \left(m + \frac{1+a^2}{1-a^2}\right) \\ &= q_1(m) \quad \text{(say)},\end{aligned}$$

$$\begin{aligned}&\sum_{j_1,j_2=-\infty}^{\infty} \gamma(j_1)\gamma(j_2)\gamma(m-j_1-j_2) \\ &\quad = \left(\frac{\sigma^2}{1-a^2}\right)^3 a^m \left\{\frac{1}{2}m(m+1) + m\frac{1+2a^2}{1-a^2} + \frac{1+4a^2+a^4}{(1-a^2)^2}\right\} \\ &\quad = q_2(m) \quad \text{(say)}\end{aligned}$$

and $\sum_{j=1}^{n} G_\theta(\lambda_j) f(\lambda) = o(1)$. It is easily seen that

$$\begin{aligned}W_{11} &= 2\left(\frac{1}{2\pi}\right)^2 \left\{\frac{1}{2}q_1(2m) - 2\theta q_1(m) + \left(\frac{1}{2} + \theta^2\right) q_1(0)\right\}, \\ W_{11,1} &= -2\left(\frac{1}{2\pi}\right)^2 \{q_1(m) - \theta q_1(0)\}, \\ W_{111} &= 6\left(\frac{1}{2\pi}\right)^3 \left\{\frac{1}{4}q_2(3m) - \frac{3}{2}\theta q_2(2m) + \left(\frac{3}{4} + 3\theta^2\right) q_2(m) - \left(\frac{3}{2}\theta + \theta^3\right) q_2(0)\right\}, \\ A_1 &= -\frac{\sigma^2}{2\pi}\frac{ma^m}{1-a^2} + o(1), \\ A_{1,1} &= -\frac{\sigma^2}{2\pi}\frac{1}{1-a^2}.\end{aligned}$$

Hence we obtain

$$\begin{aligned}m_3 &= \frac{4}{9}\frac{W_{111}(A_{1,1}h)^3}{{W_{11}}^3}, \\ m_2 &= -\frac{2}{3}\frac{W_{111}(A_{1,1}h)^3}{{W_{11}}^3} + \frac{4}{3}\frac{W_{111}A_{1,1}h}{{W_{11}}^2}, \\ m_1 &= -\frac{2}{3}\frac{W_{111}(A_{1,1}h)^3}{{W_{11}}^3} + \frac{W_{11,1}{A_{1,1}}^2h^3}{{W_{11}}^2} - 2\frac{A_1A_{1,1}h}{W_{11}}, \\ m_0 &= \frac{2}{9}\frac{W_{111}(A_{1,1}h)^3}{{W_{11}}^3} + \frac{W_{11,1}{A_{1,1}}^2h^3}{{W_{11}}^2} - \frac{2}{3}\frac{W_{111}A_{1,1}h}{{W_{11}}^2}, \\ \Delta &= \frac{(A_{1,1}h)^2}{W_{11}}.\end{aligned}$$

Consider the power function

$$P(a,\alpha) = 1 - P_{\theta_0+n^{-1/2}h}[-2\log R_n(\theta_0) < z_\alpha],$$

with $m = 1$, $\sigma^2 = 1$, $h = 1$, $n = 100$, and significance level α. Figure 1 plotted the first (solid) and the second (dashed) order approximations for the power function $P(a, 0.05)$ with $0 < a < 0.8$ given in Theorem 1. From this figure, we observe that the second order approximation is more powerful than the first order approximation.

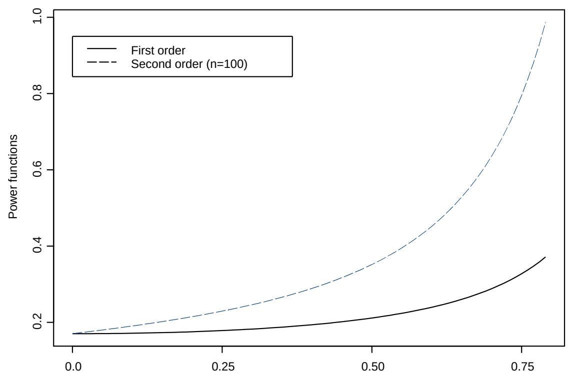

Figure 1: The power function $P(a, 0.05)$ with $0 < a < 0.8$.

Figure 2 plotted the first (solid) and the second (dashed) order approximations for the power function $P(0.3, \alpha)$ with $0 < \alpha < 0.15$. From this figure, we observe that the second order approximation is more powerful than the first order approximation.

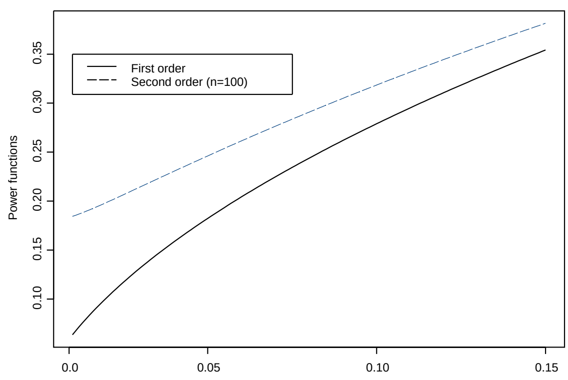

Figure 2: The power function $P(0.3, \alpha)$ with $0 < \alpha < 0.15$.

4. Proof.

Proof of Lemma 1. Expanding $g_{i,\theta_0}(\lambda)$ in a Taylor series at $\theta = \theta_1$, we get

$$g_{i,\theta_0}(\lambda) = g_{i,\theta_1}(\lambda) - \frac{1}{\sqrt{n}}\partial_a g_{i,\theta_1}(\lambda)h^a + \frac{1}{2n}\partial^2_{ab} g_{i,\theta_1}(\lambda)h^a h^b + O(n^{-3/2}). \tag{7}$$

Since

$$\begin{aligned} g_{i,\theta_0}(\lambda)[I_n(\lambda) - E\{I_n(\lambda)\}] = g_{i,\theta_1}(\lambda)[I_n(\lambda) - E\{I_n(\lambda)\}] \\ - \frac{1}{\sqrt{n}}\partial_a g_{i,\theta_1}(\lambda)h^a[I_n(\lambda) - E\{I_n(\lambda)\}] \\ + \frac{1}{2n}\partial^2_{ab} g_{i,\theta_1}(\lambda)h^a h^b[I_n(\lambda) - E\{I_n(\lambda)\}] + o_p(n^{-1}), \end{aligned}$$

$$Z_i = \tilde{Z}_i - \frac{1}{\sqrt{n}}\tilde{Z}_{i,a}h^a + o_p(n^{-1/2}). \tag{8}$$

From (4) and (7), we have

$$\begin{aligned} C_i &= \frac{1}{\sqrt{n}}\sum_{l=1}^{n} g_{i,\theta_0}(\lambda_l)\Big\{f(\lambda_l) - \frac{1}{n}b(\lambda_l)\Big\} + o(n^{-1/2}) \\ &= -\frac{1}{n}\sum_{l=1}^{n}\partial_a g_{i,\theta_1}(\lambda_l)h^a f(\lambda_l) + \frac{1}{\sqrt{n}}\sum_{l=1}^{n} g_{i,\theta_1}(\lambda_l)\Big\{f(\lambda_l) - \frac{1}{n}b(\lambda_l)\Big\} \\ &\quad + \frac{1}{2n^{3/2}}\sum_{l=1}^{n}\partial^2_{ab} g_{i,\theta_1}(\lambda_l)h^a h^b f(\lambda_l) + o(n^{-1/2}) \\ &= -\tilde{A}_{i,a}h^a + \frac{1}{\sqrt{n}}\tilde{A}_i + \frac{1}{2\sqrt{n}}\tilde{A}_{i,ab}h^a h^b + o(n^{-1/2}). \end{aligned} \tag{9}$$

Also it is easily seen that

$$\begin{aligned} V_{ij} &= \tilde{W}_{ij} - \frac{1}{\sqrt{n}}(\tilde{W}_{ij,a} + \tilde{W}_{ji,a})h^a + O(n^{-1}), \\ V_{ijk} &= \tilde{W}_{ijk} + O(n^{-1/2}). \end{aligned} \tag{10}$$

Substituting (8)-(10) into (6), after simple algebra, we obtain Lemma 1. □

Proof of Lemma 2. (i) Recalling (1) under $\theta = \theta_1$, we have

$$\begin{aligned} \mathrm{cum}[\tilde{Z}_i, \tilde{Z}_{j,a}] &= \frac{1}{n}\sum_{l_1,l_2=1}^{n} g_{i,\theta_1}(\lambda_{l_1})\partial_a g_{j,\theta_1}(\lambda_{l_2})\mathrm{cum}[I_n(\lambda_{l_1}), I_n(\lambda_{l_2})] \\ &= \frac{2}{n}\sum_{l=1}^{n} g_{i,\theta_1}(\lambda_l)\partial_a g_{j,\theta_1}(\lambda_l)f(\lambda_l)^2 + O(n^{-1}) \\ &= \tilde{W}_{ij,a} + O(n^{-1}). \end{aligned}$$

(ii) Note that

$$\begin{aligned} \mathrm{cum}[I_n(\lambda_{l_1}), I_n(\lambda_{l_2})^2] &= \mathrm{cum}[I_n(\lambda_{l_1}), I_n(\lambda_{l_2}), I_n(\lambda_{l_2})] \\ &\quad + 2\mathrm{cum}[I_n(\lambda_{l_1}), I_n(\lambda_{l_2})]\mathrm{cum}[I_n(\lambda_{l_2})]. \end{aligned}$$

From Theorem 4.3.2 in Brillinger (2001) it is seen that

$$\begin{aligned}\mathrm{cum}[\tilde{Z}_i, \tilde{Z}_{jk}] &= \frac{1}{n}\sum_{l_1,l_2=1}^{n} g_{i,\theta_1}(\lambda_{l_1})g_{j,\theta_1}(\lambda_{l_2})g_{k,\theta_1}(\lambda_{l_2})\mathrm{cum}[I_n(\lambda_{l_1}), I_n(\lambda_{l_2})^2] \\ &= \frac{8}{n}\sum_{l=1}^{n} g_{i,\theta_1}(\lambda_l)g_{j,\theta_1}(\lambda_l)g_{k,\theta_1}(\lambda_l)f(\lambda_l)^3 + O(n^{-1}) \\ &= \frac{4}{3}\tilde{W}_{ijk} + O(n^{-1}).\end{aligned}$$

□

Proof of Theorem 1. Since the actual calculation procedure is formidable, we give a sketch of the derivation. First, we evaluate the characteristic function of $-2\log R_n(\theta_0)$,

$$\psi_n(\xi, h) = E[\exp\{-2t\log R_n(\theta_0)\}], \tag{11}$$

where $t = (-1)^{1/2}\xi$. Let $D(\theta) = \{D_{ij}(\theta)\}$ be the unique lower triangular matrix with positive diagonal such that $W(\theta) = D(\theta)D(\theta)'$. We consider the transformation

$$\begin{aligned}&Y_i(\theta) = D^{ij}(\theta)Z_j(\theta), \quad Y_{i,a}(\theta) = D^{ij}(\theta)Z_{j,a}(\theta), \quad Y_{ij}(\theta) = D^{ia}(\theta)D^{jb}(\theta)Z_{ab}(\theta), \\ &B_i(\theta) = D^{ij}(\theta)A_j(\theta), \quad B_{i,a}(\theta) = D^{ij}(\theta)A_{j,a}(\theta), \quad B_{i,ab}(\theta) = D^{ij}(\theta)A_{j,ab}(\theta), \\ &U_{ij,a}(\theta) = D^{ik}(\theta)D^{jl}(\theta)W_{kl,a}(\theta), \quad U_{ijk}(\theta) = D^{ia}(\theta)D^{jb}(\theta)D^{kc}(\theta)W_{abc}(\theta),\end{aligned} \tag{12}$$

where $D^{ij}(\theta)$ is the (i,j) component of the inverse matrix of $D(\theta)$. Then the stochastic expansion of $-2\log R_n(\theta_0)$ can be rewritten as

$$\begin{aligned}-2\log R_n(\theta_0) &= (\tilde{Y}_i - \tilde{B}_{i,a}h^a)(\tilde{Y}_i - \tilde{B}_{i,b}h^b) \\ &\quad + \frac{2}{\sqrt{n}}\tilde{U}_{ij,c}h^c(\tilde{Y}_i - \tilde{B}_{i,a}h^a)(\tilde{Y}_j - \tilde{B}_{j,b}h^b) \\ &\quad + \frac{1}{\sqrt{n}}(-2\tilde{Y}_{i,b}h^b + 2\tilde{B}_i + \tilde{B}_{i,bc}h^bh^c)(\tilde{Y}_i - \tilde{B}_{i,a}h^a) \\ &\quad - \frac{1}{\sqrt{n}}\tilde{Y}_{ij}(\tilde{Y}_i - \tilde{B}_{i,a}h^a)(\tilde{Y}_j - \tilde{B}_{j,b}h^b) \\ &\quad + \frac{2}{3\sqrt{n}}\tilde{U}_{ijk}(\tilde{Y}_i - \tilde{B}_{i,a}h^a)(\tilde{Y}_j - \tilde{B}_{j,b}h^b)(\tilde{Y}_k - \tilde{B}_{k,c}h^c) \\ &\quad + o_p(n^{-1/2}).\end{aligned} \tag{13}$$

The asymptotic moments (cumulants) of $\tilde{Y}_i$ are evaluated as follows:

$$\begin{aligned}&\mathrm{cum}[\tilde{Y}_i, \tilde{Y}_j] = \delta_{ij} + o(n^{-1/2}), \\ &\mathrm{cum}[\tilde{Y}_i, \tilde{Y}_j, \tilde{Y}_k] = \frac{1}{\sqrt{n}}\Big(\frac{4}{3}\tilde{U}_{ijk} + \tilde{M}_{ijk}\Big) + o(n^{-1/2}),\end{aligned} \tag{14}$$

where

$$\begin{aligned}&M_{ijk}(\theta) = D^{ia}(\theta)D^{jb}(\theta)D^{kc}(\theta)N_{abc}(\theta), \\ &N_{ijk}(\theta) = 4\frac{1}{2\pi}\int_{-\pi}^{\pi} g_{i,\theta}(\lambda)g_{j,\theta}(\mu)g_{k,\theta}(\lambda+\mu)f_3(\lambda,\mu)f_3(-\lambda,-\mu)d\lambda d\mu,\end{aligned}$$

and $f_3(\lambda, \mu)$ is the third order cumulant spectral density of $\{X_t\}$.

We now evaluate the characteristic function (11) using the stochastic expansion (13). We take the expectation in two steps. First we consider the conditional expectation given $\tilde{Y}_1, \ldots, \tilde{Y}_p$ and then we evaluate the expectation with respect to $\tilde{Y}_1, \ldots, \tilde{Y}_p$. Relevant conditional expectations (see Takemura and Kuriki, 1996) are

$$E(\tilde{Y}_{\alpha,a}|\tilde{Y}_1, \ldots, \tilde{Y}_p) = \tilde{U}_{i\alpha,a}\tilde{Y}_i + o(1),$$
$$E(\tilde{Y}_{\alpha\beta}|\tilde{Y}_1, \ldots, \tilde{Y}_p) = \frac{4}{3}\tilde{U}_{\alpha\beta i}\tilde{Y}_i + o(1).$$

Thus we can carry out the calculation of the conditional expectation of the characteristic function.

$$\begin{aligned}\psi_n(\xi, h) = E\Big[&\exp\Big\{t(\tilde{Y}_i - \tilde{B}_{i,a}h^a)(\tilde{Y}_i - \tilde{B}_{i,b}h^b)\Big\}\\ &\times \Big\{1 + \frac{2}{\sqrt{n}}\tilde{U}_{ij,c}h^c(\tilde{Y}_i - \tilde{B}_{i,a}h^a)(\tilde{Y}_j - \tilde{B}_{j,b}h^b)\\ &+ \frac{1}{\sqrt{n}}(-2\tilde{U}_{ji,b}h^b\tilde{Y}_j + 2\tilde{B}_i + \tilde{B}_{i,bc}h^bh^c)(\tilde{Y}_i - \tilde{B}_{i,a}h^a)\\ &- \frac{4}{3\sqrt{n}}\tilde{U}_{ijk}\tilde{Y}_k(\tilde{Y}_i - \tilde{B}_{i,a}h^a)(\tilde{Y}_j - \tilde{B}_{j,b}h^b)\\ &+ \frac{2}{3\sqrt{n}}\tilde{U}_{ijk}(\tilde{Y}_i - \tilde{B}_{i,a}h^a)(\tilde{Y}_j - \tilde{B}_{j,b}h^b)(\tilde{Y}_k - \tilde{B}_{k,c}h^c)\Big\}\Big] + o(n^{-1/2}).\end{aligned}$$

From (14) and the Edgeworth expansion of $\tilde{Y}_1, \ldots, \tilde{Y}_p$ (see Section 4.1 in Taniguchi and Kakizawa, 2000), we obtain

$$\psi_n(\xi, \varepsilon) = \exp\left(\frac{t\tilde{B}_{i,a}\tilde{B}_{i,b}h^ah^b}{1-2t}\right)(1-2t)^{-p/2}\left\{1 + \frac{1}{\sqrt{n}}\sum_{j=0}^{3} m_j^*(1-2t)^{-j}\right\} + o(n^{-1/2}), \tag{15}$$

where

$$\begin{aligned}m_3^* &= \left(\frac{4}{9}\tilde{U}_{ijk} - \frac{1}{6}\tilde{M}_{ijk}\right)\tilde{B}_{i,a}\tilde{B}_{j,b}\tilde{B}_{k,c}h^ah^bh^c,\\ m_2^* &= \left(-\frac{2}{3}\tilde{U}_{ijk} + \frac{1}{2}\tilde{M}_{ijk}\right)\tilde{B}_{i,a}\tilde{B}_{j,b}\tilde{B}_{k,c}h^ah^bh^c + \left(\frac{4}{3}\tilde{U}_{ijj} - \frac{3}{2}\tilde{M}_{ijj}\right)\tilde{B}_{i,a}h^a,\\ m_1^* &= \left(-\frac{2}{3}\tilde{U}_{ijk}\tilde{B}_{i,a}\tilde{B}_{j,b} - \frac{1}{2}\tilde{M}_{ijk}\tilde{B}_{i,a}\tilde{B}_{j,b} + 2\tilde{U}_{jk,a}\tilde{B}_{j,b} - \tilde{B}_{k,ab}\right)\tilde{B}_{k,c}h^ah^bh^c\\ &\quad + (\tilde{M}_{ijj} - 2\tilde{B}_i)\tilde{B}_{i,a}h^a,\\ m_0^* &= \left(\frac{2}{9}\tilde{U}_{ijk} + \frac{1}{6}\tilde{M}_{ijk}\right)(\tilde{B}_{i,a}\tilde{B}_{j,b}\tilde{B}_{k,c}h^ah^bh^c - 3\tilde{B}_{i,a}h^a\delta_{jk}).\end{aligned}$$

We expand the right hand side of (15) in a Taylor series at $\theta = \theta_0$, and then inverting (15) by Fourier inverse transform we can prove Theorem 1. □

Acknowledgments The author would like to express his sincere thanks to Professor Masanobu Taniguchi for his encouragement and guidance.

References

[1] Brillinger, D.R. (2001). *Time series: Data analysis and theory.* Philadelphia: Society for Industrial and Applied Mathematics.

[2] Chen, S.X. (1993). On the accuracy of empirical likelihood confidence regions for linear regression model. *Ann. Inst. Statist. Math.* **45**, 621–637.

[3] Chen, S.X. (1994). Empirical likelihood confidence intervals for linear regression coefficients. *J. Multivariate Anal.* **49**, 24–40.

[4] Chen, S.X. and Cui, H. (2006). On Bartlett correction of empirical likelihood in the presence of nuisance parameters. *Biometrika* **93**, 215–220.

[5] DiCiccio, T., Hall, P. and Romano, J. (1991). Empirical likelihood is Bartlett-correctable. *Ann. Statist.* **19**, 1053–1061.

[6] Hall, P. and La Scala, B. (1990). Methodology and algorithms of empirical likelihood. *Internat. Statist. Rev. 58*, 109–127.

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

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

[9] Nordman, D.J. and Lahiri, S.N. (2006). A frequency domain empirical likelihood for short- and long-range dependence. *Ann. Statist.* **34**, 3019–3050.

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

[11] Owen, A.B. (1990). Empirical likelihood ratio confidence regions. *Ann. Statist.* **18**, 90–120.

[12] Owen, A.B. (1991). Empirical likelihood for linear models. *Ann. Statist.* **19**, 1725–1747.

[13] Qin, J. and Lawless, J. (1994). Empirical likelihood and general estimating equations. *Ann. Statist.* **22**, 300–325.

[14] Takemura, A. and Kuriki, S. (1996). A proof of independent Bartlett correctability of nested likelihood ratio tests. *Ann. Inst. Statist. Math.* **48**, 603–620.

[15] Taniguchi, M. and Kakizawa, Y. (2000). *Asymptotic theory of statistical inference for time series.* New York: Springer-Verlag.

*Department of Applied Mathematics, Waseda University, 3-4-1, Okubo, Shinjuku-ku, Tokyo, 169-8555, Japan
E-mail: k-tamaki@toki.waseda.jp