Statistica Sinica **22** (2012), 1287-1304
doi:http://dx.doi.org/10.5705/ss.2011.113

JACKKNIFED WHITTLE ESTIMATORS

Masanobu Taniguchi[1], Kenichiro Tamaki[1],
Thomas J. DiCiccio[2] and Anna Clara Monti[3]

[1] *Waseda University,* [2] *Cornell University and* [3] *University of Sannio*

Abstract: The Whittle estimator (Whittle (1962)) is widely used in time series analysis. Although it is asymptotically Gaussian and efficient, this estimator suffers from large bias, especially when the underlying process has nearly unit roots. In this paper, we apply the jackknife technique to the Whittle likelihood in the frequency domain, and we derive the asymptotic properties of the jackknifed Whittle estimator. In particular, the second-order bias of the jackknifed estimator is shown to vanish for non-Gaussian stationary processes when the unknown parameter is innovation-free. The effectiveness of the jackknife technique for reducing the bias of the Whittle estimator is demonstrated in numerical studies. Since the Whittle estimator is applicable in many fields, including the natural sciences, signal processing, and econometrics, the bias-reduced jackknifed Whittle estimator can have widespread use.

Key words and phrases: Asymptotic efficiency, innovation-free, jackknife, second-order bias, spectral density, stationary process, Whittle estimator.

1. Introduction

Quenouille (1949) introduced a technique for reducing the bias of a serial correlation estimator by splitting the sample of size n into two sub-samples. Quenouille (1956) generalized this method to splitting the sample into g groups, each of size h, so that $n = gh$. This procedure, termed the jackknife by Tukey (1958), has been employed in various estimation and testing problems. In using the jackknife method for the estimation of ratios, Rao (1965) proved that $g = n$ is the optimal choice of g for a normal auxiliary distribution. The classical jackknife was extended by Schucany, Gray, and Owen (1971) to a more general type of estimator called the generalized jackknife. A comprehensive survey of these developments was given by Miller (1974). Regarding higher-order asymptotic theory, Akahira (1989) showed that the asymptotic deficiency of the jackknife estimator relative to a bias-adjusted maximum likelihood estimator is 0.

For certain types of stochastic processes, mainly those with stationary independent increments, Gray, Watkins, and Adams (1972) introduced a counterpart of the jackknife as a sort of intensity index and gave a necessary and sufficient condition for the proposed jackknife estimator to be unbiased. Künsch

(1989) considered the jackknife in the context of estimating a functional of the m-dimensional marginal distribution of a general stationary process. Based on the jackknifed m-dimensional empirical distribution, Künsch (1989) proposed an estimator of the variance of the functional estimator and derived the asymptotic mean and variance of the variance estimator. For analyzing seismic data, Vernon (1994) proposed a jackknifed spectral estimator of window type and considered its jackknife variance estimator.

For stationary processes, the Whittle estimator $\hat{\theta}_n$ is one of the most fundamental estimators of the spectral parameter θ. Dzhaparidze (1986) showed that $\hat{\theta}_n$ is approximately the maximum likelihood estimator under Gaussian assumptions and that it is first-order asymptotically efficient. When θ is a coefficient of a Gaussian ARMAprocess, Taniguchi (1983) proved that $\hat{\theta}_n$ is second-order asymptotically efficient and calculated the second-order bias. Since the second-order bias becomes large when the process has nearly a unit root, a method for reducing bias is desirable.

In this paper, we propose a jackknifed Whittle estimator $\hat{\theta}_{\text{JK}}$ for the spectral parameter θ of a general non-Gaussian stationary processes; the jackknifing is performed in the frequency domain. We show that $\hat{\theta}_{\text{JK}}$ is asymptotically equivalent to $\hat{\theta}_n$, and we calculate the second-order biases of $\hat{\theta}_n$ and $\hat{\theta}_{\text{JK}}$. Unfortunately, both of these estimators fail to be second-order unbiased in general. To eliminate the second-order bias, we modify the Whittle likelihood, thereby obtaining a modified Whittle estimator $\hat{\theta}_n^*$ and the corresponding jackknifed version $\hat{\theta}_{\text{JK}}^*$. It is shown that $\hat{\theta}_n^*$ and $\hat{\theta}_{\text{JK}}^*$ are asymptotically equivalent and that $\hat{\theta}_{\text{JK}}^*$ is second-order unbiased if θ is innovation-free, although $\hat{\theta}_n^*$ is not unbiased to second-order. Note that the second-order bias of $\hat{\theta}_{\text{JK}}^*$ is shown to hold even if the process is non-Gaussian. The modified jackknifed variance ${\hat{V}_{\text{JK}}^*}{}^2$ is also given, and the standardized statistic $n^{1/2}(\hat{\theta}_{\text{JK}}^* - \theta)/\hat{V}_{\text{JK}}^*$ is shown to be asymptotically standard normal. These results are also generalized to the case where the unknown parameter θ is vector-valued. Numerical studies are provided confirming that the absolute bias of $\hat{\theta}_{\text{JK}}^*$ is remarkably smaller than the absolute biases of $\hat{\theta}_n$, $\hat{\theta}_{\text{JK}}$ and $\hat{\theta}_n^*$ in finite samples. These results are compelling because Whittle estimation is fundamental, and widely used in various applications.

2. Setting and Notation

Suppose that $\{X_t : t \in \mathbb{Z}\}$ is generated by

$$X_t = \sum_{l=0}^{\infty} a_l u_{t-l}, \tag{2.1}$$

where $\{a_l\}$ satisfies $\sum_{l=0}^{\infty} |a_l| < \infty$, $\{u_t\}$ is a sequence of independent and identically distributed random variables having $E(u_t) = 0$, $E(u_t^2) = \sigma^2$, and finite fourth-order cumulant for all $t \in \mathbb{Z}$, $\mathbb{Z}$ the set of all integers. To simplify the notation and discussion, we initially assume that a_l and σ^2 are functions of an unknown scalar parameter $\theta \in \Theta$, $a_l = a_l(\theta)$ and $\sigma^2 = \sigma^2(\theta)$. Here, $a_l(\theta)$ and $\sigma^2(\theta)$ are assumed to be five times continuously differentiable with respect to θ. In Section 4, the results are extended to the case where θ is vector-valued. Note that $\{X_t\}$ is second-order stationary and has spectral density function

$$f_\theta(\lambda) = \frac{\sigma^2(\theta)}{2\pi} \left| \sum_{l=0}^{\infty} a_l(\theta) e^{il\lambda} \right|^2 .$$

Let $R(s) \equiv E(X_t X_{t+s})$ be the autocovariance function of $\{X_t\}$. Based on the observed stretch $\{X_1, \ldots, X_n\}$ of $\{X_t\}$, the Whittle estimator is

$$\hat{\theta}_n = \arg\max_{\theta \in \Theta} D_n(\theta),$$

where $D_n(\theta)$ is the Whittle likelihood defined by

$$D_n(\theta) = -\frac{1}{2} \sum_{s=1}^{n} \{\log f_\theta(\lambda_s) + f_\theta(\lambda_s)^{-1} \bar{I}_n(\lambda_s)\},$$

with $\lambda_s = 2\pi s/n$ and $\bar{I}_n(\lambda) = (2\pi n)^{-1} |\sum_{t=1}^{n} X_t e^{it\lambda}|^2$. Dzhaparidze (1986) showed that, when $\{u_t\}$ is Gaussian, $D_n(\theta)$ is an approximation to the log-likelihood and that $\hat{\theta}_n$ is asymptotically optimal, i.e., first-order asymptotically efficient. For the case where θ is a coefficient of a Gaussian ARMA process, Taniguchi (1983) showed that $\hat{\theta}_n$ is second-order asymptotically efficient, and hence, a satisfactory estimator. However, if the process has a root near to 1, Taniguchi (1983, p.168) showed that $\hat{\theta}_n$ can have very large second-order bias. Consideration of a jackknifed version of the Whittle estimator is motivated by the need for bias reduction.

Calculations are simplified by using the following notation. Let $D_n^{(i)} = \partial^i D_n(\theta) / \partial\theta^i$ $(i = 1, 2, 3)$, and

$$Z_1 \equiv n^{-1/2} D_n^{(1)} = \frac{1}{n^{1/2}} \sum_{t=1}^{n} \rho_t, \quad Z_2 \equiv n^{-1/2} \left\{ D_n^{(2)} - E\left(D_n^{(2)}\right) \right\} = \frac{1}{n^{1/2}} \sum_{t=1}^{n} \gamma_t,$$

$$I_n \equiv n^{-1} E\left(-D_n^{(2)}\right) = \frac{1}{n} \sum_{t=1}^{n} \Delta_t, \quad L_n \equiv n^{-1} E\left(-D_n^{(3)}\right) = \frac{1}{n} \sum_{t=1}^{n} \kappa_t,$$

where ρ_t, γ_t, Δ_t, and κ_t are defined appropriately in terms of the addends of

$D_n(\theta)$. Furthermore, let

$$W(\lambda) \equiv \{f_\theta(\lambda)\}^{-1} \frac{\partial}{\partial\theta} f_\theta(\lambda), \quad Y(\lambda) \equiv \{f_\theta(\lambda)\}^{-2} \frac{\partial}{\partial\theta} f_\theta(\lambda),$$

$$b_\theta(\lambda) \equiv \frac{1}{2\pi} \sum_{l=-\infty}^{\infty} |l| R(l) e^{il\lambda}, \tag{2.2}$$

$$B \equiv \frac{1}{4\pi} \int_{-\pi}^{\pi} Y(\lambda) b_\theta(\lambda) d\lambda, \qquad I \equiv \frac{1}{4\pi} \int_{-\pi}^{\pi} \{W(\lambda)\}^2 d\lambda, \tag{2.3}$$

$$K \equiv \frac{1}{2\pi} \int_{-\pi}^{\pi} \{W(\lambda)\}^3 d\lambda, \qquad J \equiv -K + \frac{1}{4\pi} \int_{-\pi}^{\pi} Y(\lambda) \frac{\partial^2}{\partial\theta^2} f_\theta(\lambda) d\lambda, \tag{2.4}$$

$$I^{\text{NG}} \equiv \frac{1}{8\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} Y(\lambda) Y(\mu) Q_4(-\lambda, \mu, -\mu) d\lambda d\mu, \tag{2.5}$$

$$J^{\text{NG}} \equiv \frac{1}{8\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} Y(\lambda) \frac{\partial^2}{\partial\theta^2} \{f_\theta(\mu)^{-1}\} Q_4(-\lambda, \mu, -\mu) d\lambda d\mu, \tag{2.6}$$

where $Q_4(\lambda, \mu, \omega)$ is the fourth-order cumulant of the spectral density of $\{X_t\}$.

Noting that

$$\frac{\partial^2}{\partial\theta^2} \log f_\theta = -\, f_\theta^{-2} \left(\frac{\partial}{\partial\theta} f_\theta\right)^2 + f_\theta^{-1} \frac{\partial^2}{\partial\theta^2} f_\theta,$$
$$\frac{\partial^2}{\partial\theta^2} f_\theta^{-1} = 2 f_\theta^{-3} \left(\frac{\partial}{\partial\theta} f_\theta\right)^2 - f_\theta^{-2} \frac{\partial^2}{\partial\theta^2} f_\theta,$$

and that $E(\bar{I}_n(\lambda)) = f_\theta(\lambda) + O(n^{-1})$, we have $I_n = I + O(n^{-1})$. Also it follows from Theorem 7.6.1 of Brillinger (2001) that

$$\begin{aligned} E({Z_1}^2) &= I + I^{\text{NG}} + O(n^{-1}), \quad E(Z_1 Z_2) = J + J^{\text{NG}} + O(n^{-1}), \\ L_n &= 3J + K + O(n^{-1}). \end{aligned} \tag{2.7}$$

Evidently, if $\{X_t\}$ is Gaussian, then I^{NG} and J^{NG} are both 0. Furthermore, if θ is innovation-free, i.e., if $\sigma^2(\theta)$ does not depend on θ, then Lemma A2.2, Corollary 3.1, Proposition 3.1, and Remark 3.1 of Hosoya and Taniguchi (1982) yield

$$I^{\text{NG}} = 0, \quad J^{\text{NG}} = 0, \tag{2.8}$$

even if the process $\{X_t\}$ is non-Gaussian. The assumption that θ is innovation free is not restrictive; for example, when $\{X_t\} \sim \text{ARMA}(p, q)$ and θ is an ARMA coefficient of $\{X_t\}$, then θ is innovation-free. To ensure that $I^{\text{NG}} = 0$ and $J^{\text{NG}} = 0$ for subsequent calculations, we make the following assumption.

Assumption 1 (G-I)**.** *The process $\{X_t\}$ is Gaussian or the parameter θ is innovation-free.*

3. Jackknifed Whittle Estimators

We propose jackknifed Whittle estimators for the process (2.1) with unknown parameter θ. The advantage of working in the frequency domain is that, in the Whittle likelihood, the observations are replaced by the variables $f_\theta(\lambda_s)^{-1}\bar{I}_n(\lambda_s)$ $(s = 1, \ldots, n)$, which are asymptotically independent. Consequently, a dependent-data problem is reduced to an independent-data one, and thus the usual jackknife technique, designed for independent data, can be applied in this stochastic-process framework.

Let

$$D_{n-1}^{(j)}(\theta) \equiv -\frac{1}{2}\sum_{\substack{s=1 \\ s\neq j}}^{n}\{\log f_\theta(\lambda_s) + f_\theta(\lambda_s)^{-1}\bar{I}_n(\lambda_s)\},$$

$$\hat{\theta}_{n-1}^{(j)} \equiv \arg\max_{\theta\in\Theta} D_{n-1}^{(j)}(\theta).$$

The jackknifed Whittle estimator is

$$\hat{\theta}_{\mathrm{JK}} \equiv n\hat{\theta}_n - \frac{n-1}{n}\sum_{j=1}^{n}\hat{\theta}_{n-1}^{(j)}.$$

To derive the asymptotic properties of $\hat{\theta}_{\mathrm{JK}}$, we require the following proposition.

Proposition 1. The jackknifed Whittle estimator $\hat{\theta}_{\mathrm{JK}}$ has the stochastic expansion

$$\begin{aligned} n^{1/2}(\hat{\theta}_{\mathrm{JK}} - \theta) = &-\frac{1}{n^{3/2}}I_n^{-2}\sum_{j=1}^{n}\rho_j\gamma_j \\ &+\frac{1}{2n^{3/2}}I_n^{-3}(3J+K)\sum_{j=1}^{n}\rho_j^2 + n^{1/2}(\hat{\theta}_n - \theta) \\ &+o_p(n^{-1/2}). \end{aligned} \tag{3.1}$$

Taniguchi (1987) proved the validity of Edgeworth expansions for generalized Whittle-type estimators. In what follows, we use $\underline{E}$ to denote expectation derived from these Edgeworth expansions.

All proofs are given in the Appendix. Proposition 1 yields the following.

Theorem 1. 1. *The asymptotic distributions of* $n^{1/2}(\hat{\theta}_{\mathrm{JK}} - \theta)$ *and* $n^{1/2}(\hat{\theta}_n - \theta)$ *are the same.*

2. *The second-order biases of* $n^{1/2}(\hat{\theta}_{\text{JK}}-\theta)$ *and* $n^{1/2}(\hat{\theta}_n-\theta)$ *are*

$$\underline{E}\{n^{1/2}(\hat{\theta}_n-\theta)\} = -\frac{1}{n^{1/2}}I^{-1}B - \frac{1}{2n^{1/2}}I^{-2}(J+K) - \frac{1}{2n^{1/2}}I^{-3}(3J+K)I^{\text{NG}} + \frac{1}{n^{1/2}}I^{-2}J^{\text{NG}} + o(n^{-1/2}), \tag{3.2}$$

$$\underline{E}\{n^{1/2}(\hat{\theta}_{\text{JK}}-\theta)\} = -\frac{1}{n^{1/2}}I^{-1}B - \frac{1}{2n^{1/2}}I^{-3}(3J+K)I^{\text{NG}} + \frac{1}{n^{1/2}}I^{-2}J^{\text{NG}} + o(n^{-1/2}). \tag{3.3}$$

The jackknife eliminates the term $-\left(2n^{1/2}\right)^{-1}I^{-2}(J+K)$ from the second order bias of $n^{1/2}(\hat{\theta}_n-\theta)$. Thus, although the jackknife does not change the order of the bias, it does tend to reduce its magnitude, as is demonstrated in the simulation studies of Section 5.

Corollary 1. *Under Assumption (G-I),*

$$\underline{E}\{n^{1/2}(\hat{\theta}_n-\theta)\} = -\frac{1}{n^{1/2}}I^{-1}B - \frac{1}{2n^{1/2}}I^{-2}(J+K) + o(n^{-1/2}),$$
$$\underline{E}\{n^{1/2}(\hat{\theta}_{\text{JK}}-\theta)\} = -\frac{1}{n^{1/2}}I^{-1}B + o(n^{-1/2}).$$

It is known that the second-order bias of the usual maximum likelihood estimator $\hat{\theta}_{\text{ML}}$ of θ is $-(2n^{1/2})^{-1}I^{-2}(J+K)$; see Taniguchi (1983). The above results imply that jackknifing in the frequency domain reduces the asymptotic bias of $\hat{\theta}_{\text{ML}}$. When $\{X_t\}\sim \text{AR}(1)$ with autoregressive coefficient θ, it is seen from Taniguchi (1983) that $I^{-1}B=\theta$ and $I^{-2}(J+K)=2\theta$. Hence, from Corollary 1,

$$\underline{E}\left(\hat{\theta}_n\right) = \theta - \frac{2}{n}\theta + o(n^{-1}),$$
$$\underline{E}\left(\hat{\theta}_{\text{ML}}\right) = \theta - \frac{3}{n}\theta + o(n^{-1}),$$
$$\underline{E}\left(\hat{\theta}_{\text{JK}}\right) = \theta - \frac{1}{n}\theta + o(n^{-1}).$$

Let $\bar{\theta}_j \equiv n\hat{\theta}_n - (n-1)\hat{\theta}_{n-1}^{(j)}$. The jackknife variance estimator for $n^{1/2}\hat{\theta}_{\text{JK}}$ is

$$\hat{V}_{\text{JK}}^2 \equiv \frac{1}{n-1}\sum_{j=1}^{n}(\bar{\theta}_j - \hat{\theta}_{\text{JK}})^2.$$

Theorem 2. *Under Assumption (G-I), as* $n\to\infty$,

$$T_n \equiv n^{1/2}\frac{(\hat{\theta}_{\text{JK}}-\theta)}{\hat{V}_{\text{JK}}} \to N(0,1)$$

in distribution.

It follows from Theorem 2 that T_n can be used for testing and interval estimation of the innovation-free parameter θ even if the process $\{X_t\}$ is non-Gaussian.

Although jackknifing eliminates the second-order bias of the usual maximum likelihood estimator $\hat{\theta}_{\mathrm{ML}}$ when θ is innovation-free, the term $-n^{-1/2}I^{-1}B$ persists in the second-order bias of the jackknifed Whittle estimator $\hat{\theta}_{\mathrm{JK}}$. To eliminate this term, we modify the Whittle likelihood. Note that the periodogram is expressed as

$$\bar{I}_n(\lambda) = \frac{1}{2\pi} \sum_{l=-n+1}^{n-1} \left(1 - \frac{|l|}{n}\right) \hat{R}(l) e^{-il\lambda},$$

where

$$\hat{R}(l) = \frac{1}{n-|l|} \sum_{t=1}^{n-|l|} X_t X_{t+|l|}.$$

Since the quantity B appearing in the second-order bias stems from the part of the periodogram given by

$$H_1 \equiv \frac{1}{2\pi n} \sum_{l=-n+1}^{n-1} |l| \hat{R}(l) e^{-il\lambda},$$

we modify the periodogram as

$$\bar{I}_n^*(\lambda) = \frac{1}{2\pi} \sum_{l=-n+1}^{n-1} \hat{R}(l) e^{-il\lambda}, \tag{3.4}$$

whose Fejér transformation at λ is equal to $\bar{I}_n(\lambda)$; see Hannan (1970, p.506) Since the Fejér kernel tends to the delta function and $\bar{I}_n(\lambda) \geq 0$ a.e., we see that $\bar{I}_n^*(\lambda) \geq 0$ a.e., as $n \to \infty$. Define the modified Whittle likelihood to be

$$D_n^*(\theta) \equiv -\frac{1}{2} \sum_{s=1}^{n} \{\log f_\theta(\lambda_s) + f_\theta(\lambda_s)^{-1} \bar{I}_n^*(\lambda_s)\}.$$

Similarly, we define the modified Whittle estimator $\hat{\theta}_n$ and its jackknifed version $\hat{\theta}_{\mathrm{JK}}^*$ by

$$\hat{\theta}_n^* \equiv \arg\max_{\theta \in \Theta} D_n^*(\theta),$$

$$\hat{\theta}_{\mathrm{JK}}^* \equiv n\hat{\theta}_n^* - \frac{n-1}{n} \sum_{j=1}^{n} \hat{\theta}_{n-1}^{(j)*}.$$

Then we have

Theorem 3. 1. *The asymptotic distributions of* $n^{1/2}(\hat{\theta}^*_{\mathrm{JK}} - \theta)$ *and* $n^{1/2}(\hat{\theta}^*_n - \theta)$ *are the same.*

2. *The second-order biases of* $n^{1/2}(\hat{\theta}^*_{\mathrm{JK}} - \theta)$ *and* $n^{1/2}(\hat{\theta}^*_n - \theta)$ *are given by*

$$\begin{aligned}\underline{E}\{n^{1/2}(\hat{\theta}^*_n - \theta)\} &= -\frac{1}{2n^{1/2}} I^{-2}(J+K) \\ &\qquad - \frac{1}{2n^{1/2}} I^{-3}(3J+K)I^{\mathrm{NG}} + \frac{1}{n^{1/2}} I^{-2} J^{\mathrm{NG}} + o(n^{-1/2}), \\ \underline{E}\{n^{1/2}(\hat{\theta}^*_{\mathrm{JK}} - \theta)\} &= -\frac{1}{2n^{1/2}} I^{-3}(3J+K)I^{\mathrm{NG}} + \frac{1}{n^{1/2}} I^{-2} J^{\mathrm{NG}} + o(n^{-1/2}).\end{aligned}$$

Corollary 2. *Under Assumption* (*G-I*),

$$\begin{aligned}\underline{E}\{n^{1/2}(\hat{\theta}^*_n - \theta)\} &= -\frac{1}{2n^{1/2}}(J+K)I^{-2} + o(n^{-1/2}), \\ \underline{E}\{n^{1/2}(\hat{\theta}^*_{\mathrm{JK}} - \theta)\} &= o(n^{-1/2}).\end{aligned}$$

These results show that jackknifing the modified Whittle estimator eliminates the second-order bias in the innovation-free setting. Now let $\bar{\theta}^*_j \equiv n\hat{\theta}^*_n - (n-1)\hat{\theta}^{(j)*}_{n-1}$. The jackknife variance estimator estimator for $n^{1/2}\hat{\theta}^*_{\mathrm{JK}}$ is

$$\hat{V}^{*\,2}_{\mathrm{JK}} \equiv \frac{1}{n-1}\sum_{j=1}^{n}(\bar{\theta}^*_j - \hat{\theta}^*_{\mathrm{JK}})^2.$$

Theorem 4. *Under Assumption* (*G-I*), *as* $n \to \infty$,

$$T^*_n \equiv n^{1/2}(\hat{\theta}^*_{\mathrm{JK}} - \theta)/\hat{V}^*_{\mathrm{JK}} \to N(0,1)$$

in distribution.

4. Generalization to Vector Parameters

For simplicity and clarity, the discussion thus far has focused on the case where θ is a scalar. However, to make the results practically relevant, it is necessary to deal with the case where θ is vector-valued. This section handles the case where the process (2.1) has spectral density $f_\theta(\lambda)$ that depends on $\theta = (\theta^1, \ldots, \theta^p)' \in \Theta \subset \mathbb{R}^p$. Write $\partial_\alpha = \partial/\partial\theta^\alpha$, and in analogy to the previous definitions, let $Z^*_\alpha = n^{-1/2}\partial_\alpha D^*_n(\theta) = n^{-1/2}\sum_{t=1}^n \rho^{*,t}_\alpha$ and $Z^*_{\alpha\beta} = n^{-1/2}[\partial_\alpha\partial_\beta D^*_n(\theta) - E\{\partial_\alpha\partial_\beta D^*_n(\theta)\}] = n^{-1/2}\sum_{t=1}^n \gamma^{*,t}_{\alpha\beta}$ $(\alpha, \beta = 1, \ldots, p)$. The fundamental quantities (2.3)-(2.6) in the scalar-parameter case generalize to

$$\begin{aligned}
I_{\alpha\beta} &\equiv \frac{1}{4\pi}\int_{-\pi}^{\pi}\{\partial_\alpha \log f_\theta(\lambda)\}\{\partial_\beta \log f_\theta(\lambda)\}d\lambda,\\
J_{\alpha\beta\delta} &\equiv -\frac{1}{2\pi}\int_{-\pi}^{\pi}\{\partial_\alpha f_\theta(\lambda)\}\{\partial_\beta f_\theta(\lambda)\}\{\partial_\delta f_\theta(\lambda)\}\{f_\theta(\lambda)\}^{-3}d\lambda\\
&\quad+\frac{1}{4\pi}\int_{-\pi}^{\pi}\{\partial_\alpha f_\theta(\lambda)\}\{\partial_\beta\partial_\delta f_\theta(\lambda)\}\{f_\theta(\lambda)\}^{-2}d\lambda,\\
K_{\alpha\beta\delta} &\equiv \frac{1}{2\pi}\int_{-\pi}^{\pi}\{\partial_\alpha f_\theta(\lambda)\}\{\partial_\beta f_\theta(\lambda)\}\{\partial_\delta f_\theta(\lambda)\}\{f_\theta(\lambda)\}^{-3}d\lambda,\\
I^{\mathrm{NG}}_{\alpha\beta} &\equiv \frac{1}{8\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}[\partial_\alpha\{f_\theta(\lambda)\}^{-1}][\partial_\beta\{f_\theta(\mu)\}^{-1}]Q_4(-\lambda,\mu,-\mu)d\lambda d\mu,\\
J^{\mathrm{NG}}_{\alpha\beta} &\equiv \frac{1}{8\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}[\partial_\alpha\{f_\theta(\lambda)\}^{-1}][\partial_\beta\partial_\delta\{f_\theta(\mu)\}^{-1}]Q_4(-\lambda,\mu,-\mu)d\lambda d\mu,
\end{aligned}$$

for $\alpha,\beta,\delta = 1,\ldots,p$. Denote the αth components of $\hat\theta^*_n$ and $\hat\theta^*_{\mathrm{JK}}$ by $\hat\theta^{*,\alpha}_n$ and $\hat\theta^{*,\alpha}_{\mathrm{JK}}$, respectively, and let $I^{\alpha\beta}$ denote the (α,β)-component of the inverse of the $p\times p$ matrix $(I_{\alpha\beta})$. Proposition 2 and Theorems 5 and 6 are stated below without proofs, since their proofs are similar to the proofs of Proposition 1 and Theorems 1−4.

Proposition 2. The αth component of $\hat\theta^*_{\mathrm{JK}}$ has the stochastic expansion

$$\begin{aligned}
&n^{1/2}(\hat\theta^{*,\alpha}_{\mathrm{JK}}-\theta^\alpha)\\
&\quad= -\frac{1}{n^{3/2}}\sum_{j=1}^{n} I^{\alpha\alpha'}I^{\beta\beta'}\gamma^{*,j}_{\alpha'\beta}\rho^{*,j}_{\beta'}\\
&\qquad+\frac{1}{2n^{3/2}}\sum_{j=1}^{n} I^{\alpha\alpha'}I^{\beta\beta'}I^{\delta\delta'}(J_{\alpha'\beta\delta}+J_{\beta\delta\alpha'}+J_{\delta\alpha'\beta}+K_{\alpha'\beta\delta})\rho^{*,j}_{\beta'}\rho^{*,j}_{\delta'}\\
&\qquad+n^{1/2}(\hat\theta^{*,\alpha}_n-\theta^\alpha)+o_p(n^{-1/2}).
\end{aligned}$$

This expression and others throughout this section use the Einstein summation convention whereby summation over the range $1,\ldots,p$ is implied for any index appearing once as a subscript and once as a superscript.

Theorem 5. 1. *The asymptotic distributions of* $n^{1/2}(\hat\theta^*_{\mathrm{JK}}-\theta)$ *and* $n^{1/2}(\hat\theta^*_n-\theta)$ *are the same.*

2. *The second-order biases of* $n^{1/2}(\hat{\theta}^*_{\rm JK} - \theta)$ *and* $n^{1/2}(\hat{\theta}^*_n - \theta)$ *are given by*

$$\begin{aligned}
&\underline{E}\{n^{1/2}(\hat{\theta}^{*,\alpha}_n - \theta^\alpha)\} \\
&\quad = -\frac{1}{2n^{1/2}} I^{\alpha\alpha'} I^{\beta\delta}(J_{\alpha'\beta\delta} + J_{\beta\delta\alpha'} - J_{\delta\alpha'\beta} + K_{\alpha'\beta\delta}) \\
&\qquad - \frac{1}{2n^{1/2}} I^{\alpha\alpha'} I^{\beta\beta'} I^{\delta\delta'}(J_{\alpha'\beta\delta} + J_{\beta\delta\alpha'} - J_{\delta\alpha'\beta} + K_{\alpha'\beta\delta}) I^{\rm NG}_{\beta'\delta'} \\
&\qquad + \frac{1}{n^{1/2}} I^{\alpha\alpha'} I^{\beta\beta'} J^{\rm NG}_{\beta'\alpha'\beta} + o(n^{-1/2}),
\end{aligned}$$

$$\begin{aligned}
&\underline{E}\{n^{1/2}(\hat{\theta}^{*,\alpha}_{\rm JK} - \theta^\alpha)\} \\
&\quad = -\frac{1}{2n^{1/2}} I^{\alpha\alpha'} I^{\beta\beta'} I^{\delta\delta'}(J_{\alpha'\beta\delta} + J_{\beta\delta\alpha'} - J_{\delta\alpha'\beta} + K_{\alpha'\beta\delta}) I^{\rm NG}_{\beta'\delta'} \\
&\qquad + \frac{1}{n^{1/2}} I^{\alpha\alpha'} I^{\beta\beta'} J^{\rm NG}_{\beta'\alpha'\beta} + o(n^{-1/2}).
\end{aligned}$$

Corollary 3. *Under Assumption (G-I),*

$$\begin{aligned}
\underline{E}\{n^{1/2}(\hat{\theta}^{*,\alpha}_n - \theta^\alpha)\} &= -\frac{1}{2n^{1/2}} I^{\alpha\alpha'} I^{\beta\delta}(J_{\alpha'\beta\delta} + J_{\beta\delta\alpha'} - J_{\delta\alpha'\beta} + K_{\alpha'\beta\delta}) \\
&\quad + o(n^{-1/2}), \\
\underline{E}\{n^{1/2}(\hat{\theta}^{*,\alpha}_{\rm JK} - \theta^\alpha)\} &= o(n^{-1/2}).
\end{aligned}$$

As before, let $\bar{\theta}^*_j \equiv n\hat{\theta}^*_n - (n-1)\hat{\theta}^{(j)*}_{n-1}$, and now take

$$\Sigma^*_{\rm JK} \equiv \frac{1}{n-1}\sum_{j=1}^{n}(\bar{\theta}^*_j - \hat{\theta}^*_{\rm JK})(\bar{\theta}^*_j - \hat{\theta}^*_{\rm JK})'.$$

Theorem 6. *Under Assumption (G-I), as* $n \to \infty$*,*

$$T^*_n \equiv n^{1/2}(\Sigma^*_{\rm JK})^{-1/2}(\hat{\theta}^*_{\rm JK} - \theta) \to N_p(0, I)$$

in distribution, where $N_p(0, I)$ *is the* p*-dimensional standard normal distribution.*

As in the scalar-parameter case, these results imply that for vector parameters, if θ is innovation-free, jackknifing in the frequency domain reduces the second-order bias of the modified Whittle estimator, and the standardized statistic T^*_n has a pivotal asymptotic distribution, even if the process is non-Gaussian.

Assumption (G-I) is not restrictive, and the results have many practical applications, illustrated as follows. Even if $\{X_t\}$ is non-Gaussian, the results are useful in important problems such as prediction. Suppose, for example, that $\{X_t\} \sim$ non-Gaussian AR(p), i.e., that $X_t = \theta^1 X_{t-1} + \cdots + \theta^p X_{t-p} + u_t$. Then the best predictor of X_t is known to be $\theta^1 X_{t-1} + \cdots + \theta^p X_{t-p}$; see Fan and Yao (2003, p.118). Since $\theta^1, \ldots, \theta^p$ are "innovation-free," we can use the Jackknife-estimated predictor $\hat{\theta}^{*1}_{\text{JK}} X_{t-1} + \cdots + \hat{\theta}^{*p}_{\text{JK}} X_{t-p}$.

5. Numerical Investigations

We examined, by a Monte Carlo simulation study, the biases, variances, and coverage errors associated with five estimators, $\hat{\theta}_n$, $\hat{\theta}_{\text{JK}}$, $\hat{\theta}^*_n$, $\hat{\theta}^*_{\text{JK}}$, and the conditional Gaussian maximum likelihood estimator, for the AR(1) model

$$X_t = \theta X_{t-1} + \varepsilon_t,$$

and the ARMA(1,1) model

$$X_t = \theta X_{t-1} + \varepsilon_t + 0.25\varepsilon_{t-1},$$

where $\{\varepsilon_t\}$ is a sequence of independent and identically distributed random variables. For the simulation study, the innovations ε_t were taken to be distributed as $N(0,1)$, t_5, and $\chi^2_5 - 5$. Since the variance of ε_t is known in each case, the coefficient θ is innovation-free. Furthermore, the sample sizes were taken to be $n = 200$ and 500, and the parameter values $\theta = 0.3$, 0.6, 0.9, and 0.95 were used. No negative values θ were considered, since the absolute values of the biases, the variances, and the coverage errors of confidence intervals as functions of θ are nearly symmetric about 0.

Tables 1 and 2 show the values of $E\{n(\hat{\theta} - \theta)\}$ as estimated from 10,000 simulations for each of the five estimators $\hat{\theta}$ and each of the four values of θ. The values of $\text{var}\{n^{1/2}(\hat{\theta} - \theta)\}$ are shown in parentheses. For each situation, including the two non-Gaussian cases, the absolute bias of the jackknifed Whittle estimator is remarkably smaller than that of the Whittle estimator, and the jackknifed modified Whittle estimator has by far the absolute smallest bias of the five estimators. The variance of the jackknifed modified Whittle estimator is very similar to the variances of the other estimators. Thus, the reduction in bias is achieved without a noticeable increase in variance, so the mean squared error of the jackknifed modified Whittle estimator is smaller than that of the other estimators.

The performance of two-sided confidence intervals is evaluated by the sum of the absolute one-sided coverage errors. For a confidence interval $(\theta_{[\alpha/2]}, \theta_{[1-\alpha/2]})$, the sum of the absolute one-sided coverage errors is given by

$$|\text{pr}(\theta < \theta_{[\alpha/2]}) - \alpha/2| + |\text{pr}(\theta > \theta_{[1-\alpha/2]}) - \alpha/2|.$$

For each of the five estimators $\hat{\theta}$, and for each of the four values of θ, Table 3 shows estimates of the sums of the absolute one-sided coverage errors of nominal

Table 1. Estimates of $E\{n(\hat{\theta}-\theta)\}$ for five estimators $\hat{\theta}$ of the parameter θ in AR(1) models.

	$\theta=0.3$	0.6	0.9	0.95
Sample size $n=200$				
	Model $\varepsilon_t \sim N(0,1)$			
Conditional MLE	−0.73 (0.91)	−1.36 (0.66)	−1.81 (0.22)	−1.91 (0.13)
Whittle	−1.03 (0.91)	−1.94 (0.66)	−2.72 (0.23)	−2.95 (0.14)
Jackknifed Whittle	−0.76 (0.91)	−1.38 (0.67)	−1.84 (0.24)	−2.20 (0.14)
Modified Whittle	−0.74 (0.91)	−1.37 (0.66)	−1.84 (0.23)	−1.99 (0.14)
Jackknifed modified Whittle	−0.36 (0.92)	−0.61 (0.67)	−0.86 (0.23)	−1.42 (0.13)
	Model $\varepsilon_t \sim t_5$			
Conditional Gaussian MLE	−0.56 (0.90)	−1.20 (0.64)	−1.86 (0.22)	−1.89 (0.12)
Whittle	−0.87 (0.89)	−1.80 (0.64)	−2.79 (0.24)	−2.94 (0.13)
Jackknifed Whittle	−0.60 (0.90)	−1.28 (0.65)	−1.95 (0.24)	−2.31 (0.13)
Modified Whittle	−0.57 (0.90)	−1.20 (0.64)	−1.90 (0.23)	−1.99 (0.13)
Jackknifed modified Whittle	−0.20 (0.91)	−0.49 (0.65)	−0.97 (0.24)	−1.56 (0.13)
	Model $\varepsilon_t \sim \chi_5^2-5$			
Conditional Gaussian MLE	−0.72 (0.88)	−1.27 (0.63)	−1.73 (0.22)	−1.98 (0.13)
Whittle	−0.97 (0.88)	−1.83 (0.63)	−2.65 (0.22)	−3.00 (0.14)
Jackknifed Whittle	−0.70 (0.89)	−1.28 (0.64)	−1.78 (0.23)	−2.30 (0.14)
Modified Whittle	−0.72 (0.88)	−1.27 (0.63)	−1.74 (0.22)	−2.07 (0.14)
Jackknifed modified Whittle	−0.34 (0.89)	−0.51 (0.64)	−0.78 (0.23)	−1.54 (0.13)
Sample size $n=500$				
	Model $\varepsilon_t \sim N(0,1)$			
Conditional MLE	−0.48 (0.91)	−1.56 (0.64)	−1.89 (0.21)	−1.75 (0.11)
Whittle	−0.80 (0.91)	−2.16 (0.64)	−2.78 (0.21)	−2.71 (0.12)
Jackknifed Whittle	−0.51 (0.92)	−1.58 (0.65)	−1.90 (0.22)	−1.74 (0.12)
Modified Whittle	−0.48 (0.91)	−1.57 (0.64)	−1.89 (0.21)	−1.79 (0.12)
Jackknifed modified Whittle	−0.07 (0.92)	−0.77 (0.65)	−0.76 (0.21)	−0.66 (0.12)
	Model $\varepsilon_t \sim t_5$			
Conditional Gaussian MLE	−0.85 (0.90)	−0.96 (0.64)	−1.56 (0.20)	−1.83 (0.11)
Whittle	−1.15 (0.90)	−1.57 (0.65)	−2.45 (0.21)	−2.83 (0.12)
Jackknifed Whittle	−0.87 (0.90)	−1.00 (0.65)	−1.58 (0.21)	−1.90 (0.12)
Modified Whittle	−0.85 (0.90)	−0.98 (0.65)	−1.57 (0.20)	−1.88 (0.12)
Jackknifed modified Whittle	−0.45 (0.91)	−0.19 (0.65)	−0.45 (0.21)	−0.80 (0.12)
	Model $\varepsilon_t \sim \chi_5^2-5$			
Conditional Gaussian MLE	−0.48 (0.91)	−0.96 (0.63)	−1.77 (0.21)	−1.96 (0.11)
Whittle	−0.79 (0.90)	−1.57 (0.63)	−2.66 (0.21)	−2.90 (0.12)
Jackknifed Whittle	−0.50 (0.91)	−0.99 (0.64)	−1.78 (0.21)	−1.96 (0.12)
Modified Whittle	−0.48 (0.91)	−0.97 (0.63)	−1.77 (0.21)	−1.97 (0.11)
Jackknifed modified Whittle	−0.08 (0.91)	−0.17 (0.64)	−0.64 (0.21)	−0.87 (0.12)

MLE is the maximum likelihood estimator; simulated estimates of $\mathrm{var}\{n^{1/2}(\hat{\theta}-\theta)\}$ are shown in parentheses; estimates are based on 10,000 simulations.

Table 2. Estimates of $E\{n(\hat{\theta}-\theta)\}$ for five estimators $\hat{\theta}$ of the parameter θ in ARMA(1,1) models.

	$\theta = 0.3$	0.6	0.9	0.95
Sample size $n = 200$				
	Model $\varepsilon_t \sim N(0,1)$			
Conditional MLE	−0.58 (0.91)	−1.44 (0.64)	−1.97 (0.22)	−2.57 (0.13)
Whittle	−1.16 (0.91)	−2.33 (0.65)	−3.21 (0.24)	−4.00 (0.16)
Jackknifed Whittle	−0.88 (0.92)	−1.76 (0.66)	−2.27 (0.25)	−3.05 (0.15)
Modified Whittle	−0.50 (0.91)	−1.27 (0.65)	−1.73 (0.24)	−2.43 (0.14)
Jackknifed modified Whittle	−0.14 (0.92)	−0.57 (0.66)	−1.01 (0.23)	−1.95 (0.15)
	Model $\varepsilon_t \sim t_5$			
Conditional Gaussian MLE	−0.66 (0.88)	−1.35 (0.63)	−2.11 (0.22)	−2.58 (0.13)
Whittle	−1.26 (0.89)	−2.23 (0.66)	−3.41 (0.25)	−4.01 (0.15)
Jackknifed Whittle	−1.00 (0.90)	−1.70 (0.66)	−2.50 (0.25)	−3.16 (0.15)
Modified Whittle	−0.59 (0.89)	−1.17 (0.65)	−1.87 (0.24)	−2.46 (0.14)
Jackknifed modified Whittle	−0.24 (0.90)	−0.51 (0.66)	−1.21 (0.23)	−2.12 (0.14)
	Model $\varepsilon_t \sim \chi^2_5 - 5$			
Conditional Gaussian MLE	−0.59 (0.88)	−1.18 (0.63)	−1.97 (0.21)	−2.63 (0.13)
Whittle	−1.16 (0.88)	−2.06 (0.65)	−3.26 (0.24)	−4.06 (0.16)
Jackknifed Whittle	−0.89 (0.89)	−1.51 (0.65)	−2.33 (0.25)	−3.20 (0.15)
Modified Whittle	−0.52 (0.89)	−0.99 (0.65)	−1.77 (0.23)	−2.49 (0.14)
Jackknifed modified Whittle	−0.17 (0.90)	−0.31 (0.65)	−1.07 (0.23)	−2.11 (0.14)
Sample size $n = 500$				
	Model $\varepsilon_t \sim N(0,1)$			
Conditional MLE	−0.60 (0.90)	−1.29 (0.64)	−1.96 (0.20)	−2.06 (0.11)
Whittle	−1.18 (0.90)	−2.17 (0.64)	−3.20 (0.22)	−3.41 (0.12)
Jackknifed Whittle	−0.89 (0.90)	−1.58 (0.64)	−2.28 (0.22)	−2.36 (0.12)
Modified Whittle	−0.53 (0.90)	−1.11 (0.64)	−1.69 (0.21)	−1.82 (0.12)
Jackknifed modified Whittle	−0.14 (0.90)	−0.33 (0.65)	−0.68 (0.21)	−0.91 (0.12)
	Model $\varepsilon_t \sim t_5$			
Conditional Gaussian MLE	−0.55 (0.91)	−1.49 (0.63)	−1.95 (0.20)	−2.21 (0.12)
Whittle	−1.21 (0.92)	−2.43 (0.63)	−3.20 (0.21)	−3.48 (0.13)
Jackknifed Whittle	−0.92 (0.92)	−1.85 (0.64)	−2.30 (0.22)	−2.45 (0.13)
Modified Whittle	−0.53 (0.92)	−1.35 (0.63)	−1.68 (0.21)	−1.91 (0.12)
Jackknifed modified Whittle	−0.14 (0.92)	−0.58 (0.64)	−0.70 (0.21)	−1.04 (0.12)
	Model $\varepsilon_t \sim \chi^2_5 - 5$			
Conditional Gaussian MLE	−0.56 (0.91)	−1.39 (0.64)	−2.09 (0.21)	−2.04 (0.11)
Whittle	−1.14 (0.91)	−2.32 (0.65)	−3.33 (0.22)	−3.40 (0.13)
Jackknifed Whittle	−0.85 (0.91)	−1.74 (0.65)	−2.42 (0.22)	−2.38 (0.13)
Modified Whittle	−0.49 (0.91)	−1.23 (0.65)	−1.82 (0.21)	−1.82 (0.12)
Jackknifed modified Whittle	−0.10 (0.92)	−0.46 (0.65)	−0.82 (0.21)	−0.94 (0.12)

MLE is the maximum likelihood estimator; simulated estimates of $\mathrm{var}\{n^{1/2}(\hat{\theta}-\theta)\}$ are shown in parentheses; estimates are based on 10,000 simulations.

90% confidence intervals based on the asymptotic standard normal distribution of a standardized version of $n^{1/2}(\hat{\theta}-\theta)$. The coverage errors were estimated from 10,000 simulations and expressed as percentages in the table. When θ is nearly a unit root, the coverage errors of the jackknifed modified Whittle intervals are remarkably smaller than those of the Whittle intervals. Indeed, when $\theta = 0.9$ and $\theta = 0.95$, the intervals having the smallest coverage errors are the ones based on the jackknifed modified Whittle estimator in all three distributional cases.

6. Proofs

This section provides proofs of Proposition 1 and Theorems 1−4, which pertain to the case where θ is scalar-valued.

Lemma 1. *Suppose that $f_\theta(\lambda)$ is continuously five times differentiable with respect to θ. Then*

$$
\begin{aligned}
n^{1/2}&(\hat{\theta}^{(j)}_{n-1}-\theta)\\
&= n^{1/2}(\hat{\theta}_n-\theta)\\
&\quad + \frac{1}{n}I_n^{-2}\Delta_j Z_1 - \frac{1}{n^{1/2}}I_n^{-1}\rho_j \qquad (6.1)\\
&\quad - \frac{1}{n}I_n^{-2}\rho_j Z_2 - \frac{1}{n}I_n^{-2}\gamma_j Z_1 + \frac{2}{n^{3/2}}I_n^{-3}\Delta_j Z_1 Z_2 \qquad (6.2)\\
&\quad + \frac{1}{n^{3/2}}I_n^{-2}\rho_j\gamma_j\\
&\quad + \frac{1}{2n^{3/2}}I_n^{-3}\kappa_j Z_1^2 - \frac{3}{2n^{3/2}}I_n^{-4}(3J+K)\Delta_j Z_1^2 + \frac{1}{n}I_n^{-3}(3J+K)\rho_j Z_1 \qquad (6.3)\\
&\quad - \frac{1}{2n^{3/2}}I_n^{-3}(3J+K)\rho_j^2 + o_p(n^{-3/2}).
\end{aligned}
$$

Proof. For simplicity, write $D^{(j)}_{n-1}(\theta)$ and $\hat{\theta}^{(j)}_{n-1}$ as $\tilde{D}$ and $\tilde{\theta}$, respectively. Expanding the right hand side of the equation

$$0 = \frac{1}{n^{1/2}}\frac{\partial}{\partial\theta}\tilde{D}(\tilde{\theta})$$

Table 3. Estimates of sum of absolute one-sided coverage errors for nominal 90% confidence intervals for θ in AR(1) and ARMA(1,1) models.

	AR(1)				ARMA(1,1)			
	$\theta = 0.3$	0.6	0.9	0.95	0.3	0.6	0.9	0.95
Sample size $n = 200$								
	Model $\varepsilon_t \sim N(0,1)$							
Conditional MLE	0.89	1.27	3.19	4.34	0.42	1.53	2.95	7.00
Whittle	1.44	2.50	6.24	8.64	1.38	3.11	7.39	13.99
Jackknifed Whittle	1.02	1.24	2.94	4.62	0.95	2.06	3.77	9.47
Modified Whittle	0.88	1.10	3.23	4.29	0.27	1.21	2.00	7.58
Jackknifed modified Whittle	0.32	1.23	2.61	3.23	0.22	0.72	1.85	6.53
	Model $\varepsilon_t \sim t_5$							
Conditional Gaussian MLE	0.16	1.24	3.27	3.82	0.52	1.29	3.26	7.60
Whittle	0.59	2.07	6.21	8.63	1.47	3.07	7.58	13.97
Jackknifed Whittle	0.34	1.18	3.20	4.86	1.03	2.02	4.30	9.50
Modified Whittle	0.16	1.17	3.11	4.37	0.35	1.02	2.25	7.68
Jackknifed modified Whittle	0.41	0.26	2.55	3.00	0.19	0.39	1.77	6.16
	Model $\varepsilon_t \sim \chi^2_5 - 5$							
Conditional Gaussian MLE	0.58	1.06	2.73	4.81	0.20	1.14	3.16	7.39
Whittle	0.96	1.95	5.54	9.19	0.44	2.58	7.27	14.48
Jackknifed Whittle	0.50	0.91	2.31	5.00	0.10	1.62	3.97	9.91
Modified Whittle	0.50	0.77	2.71	5.33	0.31	0.61	2.40	7.28
Jackknifed modified Whittle	0.17	0.59	2.05	3.63	0.66	0.76	1.94	5.99
Sample size $n = 500$								
	Model $\varepsilon_t \sim N(0,1)$							
Conditional MLE	0.23	1.36	2.63	2.10	0.12	0.81	2.75	2.97
Whittle	0.56	2.00	4.30	5.45	0.51	1.88	5.18	7.21
Jackknifed Whittle	0.27	1.44	2.46	2.35	0.21	1.29	3.17	3.66
Modified Whittle	0.19	1.33	2.45	2.93	0.10	0.70	1.63	2.28
Jackknifed modified Whittle	0.14	0.51	1.84	1.40	0.58	0.31	1.38	2.22
	Model $\varepsilon_t \sim t_5$							
Conditional Gaussian MLE	0.62	0.62	1.55	2.81	0.21	1.16	2.63	3.56
Whittle	0.88	1.35	3.90	5.43	0.57	2.29	5.00	7.24
Jackknifed Whittle	0.58	0.69	1.82	2.25	0.28	1.63	3.01	4.23
Modified Whittle	0.49	0.76	1.85	2.49	0.07	1.16	1.68	2.30
Jackknifed modified Whittle	0.15	0.27	0.90	2.03	0.37	0.21	1.25	2.01
	Model $\varepsilon_t \sim \chi^2_5 - 5$							
Conditional Gaussian MLE	0.37	0.54	1.93	2.94	0.25	0.83	2.45	2.88
Whittle	0.47	1.21	3.80	5.86	0.47	1.96	5.09	6.97
Jackknifed Whittle	0.41	0.66	1.65	2.68	0.13	1.24	3.27	3.82
Modified Whittle	0.35	0.61	1.82	2.90	0.27	0.53	1.96	2.33
Jackknifed modified Whittle	0.45	0.36	0.98	1.96	0.68	0.41	1.27	1.95

Table entries are percentages; estimates are based on 10,000 simulations.

yields

$$
\begin{aligned}
0 =& \frac{1}{n^{1/2}}\left\{\frac{\partial}{\partial\theta}\tilde{D}(\theta)+\frac{\partial^2}{\partial\theta^2}\tilde{D}(\theta)(\tilde{\theta}-\theta)+\frac{1}{2}\frac{\partial^3}{\partial\theta^3}\tilde{D}(\theta)(\tilde{\theta}-\theta)^2+\cdots\right\}\\
=& \frac{1}{n^{1/2}}\frac{\partial}{\partial\theta}\tilde{D}(\theta)+\left\{\frac{1}{n}\frac{\partial^2}{\partial\theta^2}\tilde{D}(\theta)\right\}\sqrt{n}(\tilde{\theta}-\theta)\\
&+\frac{1}{2n^{1/2}}\left\{\frac{1}{n}\frac{\partial^3}{\partial\theta^3}\tilde{D}(\theta)\right\}\{n^{1/2}(\tilde{\theta}-\theta)\}^2+\cdots\\
=& \tilde{Z}_1+E\left\{\frac{1}{n}\frac{\partial^2}{\partial\theta^2}\tilde{D}(\theta)\right\}n^{1/2}(\tilde{\theta}-\theta)+\frac{1}{n^{1/2}}\tilde{Z}_2 n^{1/2}(\tilde{\theta}-\theta)\\
&+\frac{1}{2n^{1/2}}E\left\{\frac{1}{n}\frac{\partial^3}{\partial\theta^3}\tilde{D}(\theta)\right\}\{n^{1/2}(\tilde{\theta}-\theta)\}^2+\cdots, \qquad (6.4)
\end{aligned}
$$

where

$$
\tilde{Z}_1=\frac{1}{n^{1/2}}\frac{\partial}{\partial\theta}\tilde{D}(\theta), \quad \tilde{Z}_2=\frac{1}{n^{1/2}}\left[\frac{\partial^2}{\partial\theta^2}\tilde{D}(\theta)-E\left\{\frac{\partial^2}{\partial\theta^2}\tilde{D}(\theta)\right\}\right].
$$

The desired stochastic expansion is obtained by substituting $n^{1/2}(\tilde{\theta}-\theta)=A_1+n^{-1/2}A_2+n^{-1}A_3+\cdots$ into (6.4), and noting that $\tilde{Z}_1=Z_1-n^{-1/2}\rho_j$, $\tilde{Z}_2=Z_2-n^{-1/2}\gamma_j$, and

$$
E\left\{-\frac{1}{n}\frac{\partial^3}{\partial\theta^3}\tilde{D}(\theta)\right\}=L_n-\frac{\kappa_j}{n};
$$

see Taniguchi and Kakizawa (2000, Chap. 4) for related calculations.

Proof of Proposition 1. Note that

$$
n^{1/2}(\hat{\theta}_{\text{JK}}-\theta)=\sum_{j=1}^{n}\{n^{1/2}(\hat{\theta}_n-\theta)-n^{1/2}(\hat{\theta}_{n-1}^{(j)}-\theta)\}+\frac{1}{n}\sum_{j=1}^{n}n^{1/2}(\hat{\theta}_{n-1}^{(j)}-\theta). \qquad (6.5)
$$

It is easily seen that

$$
\sum_{j=1}^{n}(6.1)=0, \quad \sum_{j=1}^{n}(6.2)=0, \quad \sum_{j=1}^{n}(6.3)=0.
$$

By substituting the stochastic expansion of $n^{1/2}(\hat{\theta}_{n-1}^{(j)}-\theta)$ from Lemma 1 into (6.5), we obtain (3.1).

Proof of Theorem 1.

1. This result follows from (3.1).
2. Formula (3.2) follows from (2.7), and Theorem 5 and equation (5.5) of Taniguchi (1983). From (3.1) and (2.7), we obtain (3.3).

Proof of Theorem 2. From the definition of $\hat{V}^2_{\text{JK}}$, we have

$$\hat{V}^2_{\text{JK}} = \frac{n-1}{n} \sum_{j=1}^{n} \left\{ n^{1/2}(\hat{\theta}^{(j)}_{n-1} - \theta) - \frac{1}{n} \sum_{k=1}^{n} n^{1/2}(\hat{\theta}^{(k)}_{n-1} - \theta) \right\}^2 . \tag{6.6}$$

By substituting the stochastic expansion from Lemma 1 into (6.6), we obtain

$$\begin{aligned} \hat{V}^2_{\text{JK}} &= \sum_{j=1}^{n} \left(\frac{1}{n^{1/2}} I_n^{-1} \rho_j \right)^2 + \text{lower order terms} \\ &\to I^{-1}, \text{ in distribution, as } n \to \infty, \end{aligned}$$

which, together with Slutsky's lemma, produces the desired conclusion.

Proof of Theorems 3 and 4. From the definitions of $b_\theta(\lambda)$ and $\bar{I}^*_n(\lambda)$ given by (2.2) and (3.4), respectively, it follows that

$$E\{\bar{I}_n(\lambda)\} - E\{\bar{I}^*_n(\lambda)\} = -\frac{1}{n} b_\theta(\lambda) + o(n^{-1}).$$

Let

$$Z^*_1 = \frac{1}{n^{1/2}} \frac{\partial}{\partial\theta} D^*_n(\theta), \quad Z^*_2 = \frac{1}{n^{1/2}} \left[\frac{\partial^2}{\partial\theta^2} D^*_n(\theta) - E\left\{ \frac{\partial^2}{\partial\theta^2} D^*_n(\theta) \right\} \right].$$

We observe that $E(Z^*_1) = o(n^{-1/2})$, and that Z^*_1 and Z^*_2 have the same asymptotic properties as those of Z_1 and Z_2 shown in (2.7); see the proof of Lemma A3.3 of Hosoya and Taniguchi (1982). Hence, Theorems 3 and 4 follow by applying the same arguments as those used in the proofs of Theorems 1 and 2.

Acknowledgement

This paper was presented at the Fourth Brussels-Waseda Seminar on "Time Series and Financial Statistics" at ENSAI (Rennes) in 2009. The first author thanks all the attendants for their comments. This work was partially supported by Japan-Belgium Research Cooperative Program (JSPS & FNRS) and Grand-in-Aid (A) (19204009) of Japan. The authors would like to thank Professor Peter Hall and two anonymous referees for their helpful comments and suggestions, which improved the original version of this paper.

References

Akahira, M. (1989). Behaviour of jackknife estimators in terms of asymptotic deficiency under true and assumed models. *J. Japan Statist. Soc.* **19**, 179-196.

Brillinger, D. R. (2001). *Time Series: Data Analysis and Theory.* Society for Industrial and Applied Mathematics, Philadelphia, PA.

Dzhaparidze, K. (1986). *Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series.* Springer-Verlag, New York.

Fan, J. and Yao, Q. (2003). *Nonlinear Time Series.* Springer-Verlag, New York.

Gray, H. L., Watkins, T. A. and Adams, J. E. (1972). On the jackknife statistic, its extensions, and its relation to e_n-transformations. *Ann. Math. Statist.* **43**, 1-30.

Hannan, E. J. (1970). *Multiple Time Series.* John Wiley and Sons, New York.

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.

Künsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *Ann. Statist.* **17**, 1217-1241.

Miller, R. G. (1974). The jackknife—a review. *Biometrika* **61**, 1-15.

Quenouille, M. H. (1949). Approximate tests of correlation in time-series. *J. Roy. Statist. Soc. Ser. B* **11**, 68-84.

Quenouille, M. H. (1956). Notes on bias in estimation. *Biometrika* **43**, 353-360.

Rao, J. N. K. (1965). A note on estimation of ratios by Quenouille's method. *Biometrika* **52**, 647-649.

Schucany, W. R., Gray, H. L. and Owen, D. B. (1971). On bias reduction in estimation. *J. Amer. Statist. Assoc.* **66**, 524-533.

Taniguchi, M. (1983). On the second order asymptotic efficiency of estimators of Gaussian ARMA processes. *Ann. Statist.* **11**, 157-169.

Taniguchi, M. (1987). Validity of Edgeworth expansions of minimum contrast estimators for Gaussian ARMA processes. *J. Multivariate Anal.* **21**, 1-28.

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

Tukey, J. W. (1958). Bias and confidence in not-quite large samples. *Ann. Math. Statist.* **29**, 614.

Vernon, F. L. (1994). Jackknifed multiple-window spectra and coherence applied to seismicdata. *IEEE International Conference on Acoustics, Speech, and Signal Processing, Proceedings-94* **6**, 93-96.

Whittle, P. (1962). Gaussian estimation in stationary time series. *Bull. Inst. Internat. Statist.* **39**, 105-129.

Department of Applied Mathematics, Waseda University, 3-4-1, Okubo, Shinjuku-ku, Tokyo, 169-8555, Japan.

E-mail: taniguchi@waseda.jp

Faculty of Political Science and Economics, Waseda University, 1-6-1 Nishiwaseda, Shinjuku-ku, Tokyo, 169-8050, Japan.

E-mail: tamaki@waseda.jp

Department of Social Statistics, Cornell University, Ithaca, New York 14853, U.S.A.

E-mail: tjd9@cornell.edu

Faculty of Economics, University of Sannio, Via Calandra 1, 82100 Benevento, Italy.

E-mail: acmonti@unisannio.it

(Received May 2011; accepted September 2011)