Journal of the Korean Statistical Society (1999), **28**: *4, pp 503–513*

ARMA Model Identification Using the Bayes Factor†

Young Sook Son[1]

Abstract

The Bayes factor for the identification of stationary $ARMA(p,q)$ models is exactly computed using the Monte Carlo method. As priors are used the uniform prior for $(\underline{\phi}_p, \underline{\theta}_q)$ in its stationarity-invertibility region, the Jefferys prior and the reference prior that are noninformative improper for $(\mu, \sigma_\varepsilon)$.

Keywords: Stationary ARMA model; Stationarity-invertibility region; Noninformative improper prior; Jefferys prior; Reference prior; Bayes factor, Posterior probability.

1. INTRODUCTION

Suppose that $Z_1, Z_2, \ldots, Z_n$ follow a stationary $ARMA(p,q)$ model,

$$\Phi_p(B)(Z_t - \mu) = \Theta_q(B)\varepsilon_t ,$$

where $\{\varepsilon_t\}$ is a $N(0, \sigma_\varepsilon^2)$ white noise, $\Phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$, $\Theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$, and μ, σ_ε^2, $\underline{\phi}_p = (\phi_1, \phi_2, \ldots, \phi_p)$, and $\underline{\theta}_q = (\theta_1, \theta_2, \ldots, \theta_q)$ are all unknown parameters. For the stationarity and invertibility of ARMA process $(\underline{\phi}_p, \underline{\theta}_q)$ must lie in the region $C_p \times C_q$, where

$$C_p \times C_q = \{(\underline{\phi}_p, \underline{\theta}_q) : \Phi_p(x) = 0,\ |x| > 1 \text{ and } \Theta_q(y) = 0,\ |y| > 1\}.$$

The first step in the analysis of Box and Jenkins' stationary $ARMA(p,q)$ models is to determine the autoregressive order, p, and the moving average order, q, for model identification. There are several identification methods of $ARMA$ models(see Choi(1992) for them). But our interest in this paper is only on its Bayesian procedure. As a Bayesian method for $ARMA(p,q)$ model identification,

†This paper was supported by Chonnam National University Research Fund, 1997.
[1]Department of Statistics, Chonnam National University, Kwangju, 500-757, Korea.

there is a famous Schwarz(1978)'s SBC(Schwarz's Bayesian Criterion) criterion based on Laplace's asymptotic approximation. In this paper we will introduce other Bayesian criteria for $ARMA$ model identification. To do this we need to review the SBC criterion.

The marginal or the predictive density of $\underline{Z} = (Z_1, Z_2, \ldots, Z_n)'$ under the $ARMA(p,q)$ model is defined by

$$m_{(p,q)}(\underline{Z}) = \int L_{(p,q)}(\underline{\delta}|\underline{Z}) \cdot \pi_{(p,q)}(\underline{\delta}) d\underline{\delta} \ , \tag{1.1}$$

where $\underline{\delta} = (\mu, \sigma_\varepsilon, \underline{\phi}_p, \underline{\theta}_q)$, and $L_{(p,q)}(\underline{\delta}|\underline{Z})$ and $\pi_{(p,q)}(\underline{\delta})$ are a likelihood function and a prior density under the $ARMA(p,q)$ model, respectively .

As a result of Laplace's asymptotic approximation the marginal density of $\underline{Z}$ is approximated as

$$m_{(p,q)}(\underline{Z}) \approx \{L_{(p,q)}(\hat{\underline{\delta}}|\underline{Z})|\hat{I}_{(p,q)}|^{-\frac{1}{2}}\} \cdot \{(2\pi)^{\frac{p+q+2}{2}} \pi_{(p,q)}(\hat{\underline{\delta}})\} \ , \tag{1.2}$$

where $\hat{I}_{(p,q)}$ and $\hat{\underline{\delta}}$ are the observed information matrix and the maximum likelihood estimate(m.l.e.) of $\underline{\delta}$ under the $ARMA(p,q)$ model, respectively. Ignoring the term in the second brace of (1.2) and applying $-2\ln$ function, the SBC criterion is finally obtained by

$$SBC(p,q) = n \cdot \ln \hat{\sigma}^2_{\varepsilon(p,q)} + (p+q+2) \cdot \ln n \ , \tag{1.3}$$

where $\hat{\sigma}^2_{\varepsilon(p,q)}$ is the m.l.e. of σ^2_ε under the $ARMA(p,q)$ model.

Basic idea of model identification by the SBC criterion is to find p and q to maximize the marginal density, $m_{(p,q)}(\underline{Z})$, of (1.1), i.e., to minimize $SBC(p,q)$ of (1.3). Some problems are often pointed out in Laplace's approximation to the SBC criterion. First, the second brace term of (1.2) can be a dominant function for small n. Second, for nested models the SBC criterion to drop the second brace term select more complex model.(Berger and Pericchi(1994)(1996)) Third, Laplace approximation starts from the nice condition that $\ln L(\underline{\delta}|\underline{Z})$ is a smooth, bounded, and unimodal function with a maximum at the m.l.e., $\hat{\underline{\delta}}$. But, practically the maximum likelihood procedure sometimes fails to converge.

Our goal in this paper is not to drop any term in the marginal density of (1.1) but to compute exactly. Here, we use the uniform prior for $(\underline{\phi}, \underline{\theta})$ in its stationarity-invertibility region and the default Bayes priors, the Jefferys prior and the reference prior, that are noninformative improper for $(\mu, \sigma_\varepsilon)$. Then p and q are selected by posterior probabilities computed using Bayes factors.

The integral in the stationarity-invertibility region, $C_p \times C_q$, is computed by Monte Carlo method using Jones(1987)' algorithm for obtaining the ARMA parameters, $\underline{\phi}_p$ and $\underline{\theta}_q$, uniformly from $C_p \times C_q$. Also to compute the inverse of the covariance matrix of $ARMA(p,q)$ process the closed form by Leeuw(1994) is explicitly used.

Monahan(1983) dealt with the problem of model selection in a fully Bayesian analysis of stationary ARMA models. He adopted as prior distributions the uniform prior distribution for $(\underline{\phi}_p, \underline{\theta}_q)$ in its stationarity-invertibility region and the standard normal-inverse gamma conjugate prior for $(\mu, \sigma_\varepsilon)$. He proposed the numerical integration method after transforming parameters to compute the integral in the stationarity-invertibility region . But since the transformation method is many to one, the computation of the integral is very tedious and not automatic.

Varshavsky(1995) used the arithmetic intrinsic Bayes factor of Berger and Pericchi(1996) to determine the order of trend and the autoregressive order in the nonstationary autoregressive model with a trend. She assumed the uniform prior for $\underline{\phi}_p$ in its stationary region, the Jefferys prior and the reference prior that are noninformative improper for $(\mu, \sigma_\varepsilon)$ and the regression coefficient $\underline{\beta}$ in a trend. She computed the integral in the stationarity region, C_p, by Monte Carlo method using Jones(1987)' algorithm.

In the next section, the Bayes factor for the identification of $ARMA(p,q)$ models is constructed. In section 3, the procedure for computing the Bayes factor defined in section 2 is introduced. In section 4, the method discussed in this paper is applied to some time series data in Box and Jenkins(1978)' and Wei(1991)'s texts.

2. THE BAYES FACTOR FOR THE IDENTIFICATION OF ARMA MODEL

The likelihood function of $\underline{\delta} = (\mu, \sigma_\varepsilon, \underline{\phi}_p, \underline{\theta}_q)$ under the $ARMA(p,q)$ model is given by

$$L(\mu, \sigma_\varepsilon, \underline{\phi}_p, \underline{\theta}_q) = (2\pi\sigma_\varepsilon^2)^{-\frac{n}{2}} |V_{(p,q)}^{-1}|^{\frac{1}{2}} \exp\{-\frac{1}{2\sigma_\varepsilon^2}(\underline{Z} - \mu\underline{1})' V_{(p,q)}^{-1} (\underline{Z} - \mu\underline{1})\}, \quad (2.1)$$

where $\underline{1}$ is an $n \times 1$ one vector and $V_{(p,q)}$ is an $n \times n$ matrix composed of only $(\underline{\phi}_p, \underline{\theta}_q)$ such that $Cov(\underline{Z}) = \sigma_\varepsilon^2 V_{(p,q)}$.

We will consider default noninformative priors of the form,

$$\pi(\mu, \sigma_\varepsilon, \underline{\phi}_p, \underline{\theta}_q) = \pi^N(\mu, \sigma_\varepsilon)\pi(\underline{\phi}_p, \underline{\theta}_q) \ , \tag{2.2}$$

where

$$\pi(\underline{\phi}_p, \underline{\theta}_q) = \frac{I_{C_p \times C_q}(\underline{\phi}_p, \underline{\theta}_q)}{Volume(C_p \times C_q)} \ ,$$

with $I_{C_p \times C_q}(\underline{\phi}_p, \underline{\theta}_q) = 1$, if $(\underline{\phi}_p, \underline{\theta}_q) \in C_p \times C_q$ and 0, otherwise, and

$$\pi^N(\mu, \sigma_\varepsilon) \propto \frac{1}{\sigma_\varepsilon^k}, \qquad -\infty < \mu < \infty, \ \ 0 < \sigma_\varepsilon < \infty$$

with default choices of k being $k = 1$ in case of the reference prior and $k = 2$ in case of Jefferys prior. Here $\pi^N(\cdot)$ denotes the noninformative prior. Now, from (1.1), (2.1), and (2.2) the marginal density of $\underline{Z}$ under the $ARMA(p, q)$ model is defined by

$$m^N_{(p,q)}(\underline{Z}) = \int_{C_p \times C_q} \int_0^\infty \int_{-\infty}^\infty \quad L(\mu, \sigma_\varepsilon, \underline{\phi}_p, \underline{\theta}_q \mid \underline{Z}) \cdot \pi^N(\mu, \sigma_\varepsilon)\pi(\underline{\phi}_p, \underline{\theta}_q) d\mu \ d\sigma_\varepsilon d\underline{\phi}_p d\underline{\theta}_q \ . \tag{2.3}$$

Thus, after integrating over μ and σ_ε in (2.3), the Bayes factor of the $ARMA$ (p, q) model to the $ARMA(p', q')$ is obtained as

$$B^N{}_{(p,q)(p',q')}(\underline{Z}) = \frac{m^N{}_{(p,q)}(\underline{Z})}{m^N{}_{(p',q')}(\underline{Z})}$$

$$= \frac{Volume(C_p \times C_q)^{-1}}{Volume(C_{p'} \times C_{q'})^{-1}} \cdot \frac{\int_{C_p \times C_q} |V^{-1}_{(p,q)}|^{\frac{1}{2}} (\underline{1}' V^{-1}_{(p,q)} \underline{1})^{-\frac{1}{2}} R_{(p,q)}^{-\frac{1}{2}(n+k-2)} d\underline{\phi}_p d\underline{\theta}_q}{\int_{C_{p'} \times C_{q'}} |V^{-1}_{(p',q')}|^{\frac{1}{2}} (\underline{1}' V^{-1}_{(p',q')} \underline{1})^{-\frac{1}{2}} R_{(p',q')}^{-\frac{1}{2}(n+k-2)} d\underline{\phi}_{p'} d\underline{\theta}_{q'}}, \tag{2.4}$$

where

$$R_{(p,q)} = \underline{Z}' V^{-1}_{(p,q)} \underline{Z} - \frac{(\underline{1}' V^{-1}_{(p,q)} \underline{Z})^2}{\underline{1}' V^{-1}_{(p,q)} \underline{1}}.$$

Default or automatic priors such as the Jefferys prior or the reference prior are used when subjective priors for each model are not feasible. But they are often noninformative improper priors. The main difficulty in developing the default Bayes factor using the default improper prior is to be unable to determinc an unknown arbitrary multiplicative constant of improper prior. So, an improper prior, $\pi^N(\mu, \sigma_\varepsilon)$, of (2.2) can not be directly used in (2.3). But, Jefferys(1961) discussed that arbitrary multiplicative constants for the improper priors would be cancelled in Bayes factor in case of using noninformative improper priors for common parameters in the models. Our problem is just applied to this case.

The Bayes factor reflects the support of two models by data. That is, if the value of the Bayes factor $B_{(p,q)(p',q')}$ is larger than 1, it implies that the $ARMA(p,q)$ model is more supported than the $ARMA(p',q')$ model by given data.

Bayesian procedure of selecting a model among more than two models is to choose the model that gives the maximum posterior probability $P(ARMA(p,q)|\underline{Z})$ with the prior probability, $p_{(p,q)}$, of each model $ARMA(p,q)$ being true, where

$$P(ARMA(p,q) \mid \underline{Z}) = \{ \sum_{(p',q')} \frac{p_{(p',q')}}{p_{(p,q)}} B_{(p',q')(p,q)} \}^{-1}. \tag{2.5}$$

3. COMPUTATION

To compute (2.4) is reduced to solving two problems; the first is to get the inverse matrix, $V^{-1}_{(p,q)}$, satisfying $Cov(\underline{Z}) = \sigma^2_\varepsilon V_{(p,q)}$ and the second is to compute the integrals over ϕ_p andθ_q. The $V^{-1}_{(p,q)}$ matrix can be explicitly expressed by only $\underline{\phi}_p$ and $\underline{\theta}_q$ using the result of Leeuw(1994). Then the computation of the integrals in (2.4) is equivalent to the problem of computing the integral of the form

$$\int_{C_p \times C_q} \frac{g(\underline{\phi}_p, \underline{\theta}_q)}{Volume(C_p \times C_q)} d\underline{\phi}_p d\underline{\theta}_q \ , \tag{3.1}$$

where $g(\underline{\phi}_p, \underline{\theta}_q)$ is a function of only $\underline{\phi}_p$ and $\underline{\theta}_q$.

To compute (3.1) we use the method by Monahan(1984) and Jones(1987) of randomly choosing parameters from the stationarity-invertibility region of ARMA process.

There is one to one transformation between $\underline{\phi}_p = (\phi_1, \phi_2, \ldots, \phi_p)$ and the partial autocorrelation $\underline{\gamma}_p = (\gamma_1, \gamma_2, \ldots, \gamma_p)$ that maps C_p onto $[-1,1]^p$. Let $\underline{y}^{(k)} = (y_1^{(k)}, y_2^{(k)}, \ldots, y_k^{(k)})$, $k = 1, 2, \ldots, p$. Then $y_i^{(k)}$ is calculated from the recursive relation,

$$y_i^{(k)} = y_i^{(k-1)} - \gamma_k y_{k-i}^{(k-1)}, \qquad i = 1, 2, \ldots, k-1, \tag{3.2}$$

with $y_1^{(1)} = \gamma_1$ as the initial setting and $y_k^{(k)} = \gamma_k$ as the final setting. Finally, set $\underline{\phi}_p = \underline{y}^{(p)}$. For example of $p = 3$, $\phi_1 = \gamma_1 - \gamma_1\gamma_2 - \gamma_2\gamma_3$, $\phi_2 = \gamma_2 - \gamma_1\gamma_3 + \gamma_1\gamma_2\gamma_3$, and $\phi_3 = \gamma_3$. The Jacobian of the transformation is

$$\begin{aligned} J_p(\underline{\gamma}_p) &= \prod_{k=2}^{p}(1-\gamma_k)^{[\frac{k}{2}]}(1+\gamma_k)^{[\frac{1}{2}(k-1)]} \\ &= \prod_{k=2}^{p} B_{\gamma_k}\left([\tfrac{1}{2}(k+1)], [\tfrac{1}{2}k]+1\right), \end{aligned}$$

where $B_{\gamma_k}(\alpha_1, \alpha_2) = (1+\gamma_k)^{\alpha_1 - 1}(1-\gamma_k)^{\alpha_2 - 1}$ is a kernel of the rescaled beta density of γ_k defined on $[-1, 1]$ with parameters α_1 and α_2 . The relation between $\underline{\theta}_q = (\theta_1, \theta_2, \ldots, \theta_q)$ and the partial autocorrelation $\underline{\gamma}_q = (\gamma_1, \gamma_2, \ldots, \gamma_q)$ that maps C_q onto $[-1, 1]^q$ is similarly developed as that of $\underline{\phi}_p$ and $\underline{\gamma}_p$. Thus a numerical calculation of the integration in (3.1) can be done through

$$\int_{[-1,1]^{p+q}} g(\underline{\gamma}_p, \underline{\gamma}_q) \frac{J_p(\underline{\gamma}_p) J_q(\underline{\gamma}_q)}{Volume(C_p \times C_q)} d\underline{\gamma}_p d\underline{\gamma}_q ,$$

where $Volume(C_p \times C_q)^{-1}$ is a product of normalizing constants of the rescaled beta densities defined on $[-1, 1]$ with parameters $([\frac{1}{2}(k+1)], [\frac{1}{2}k]+1)$, $k = 2, 3, \ldots, p$, $k = 2, 3, \ldots, q$.

Now, the Monte Carlo method for computing the integral of (3.1) yields the following algorithm.

STEP 1 : Generate γ_k's ($k = 1, 2, \ldots, p$) independently from a rescaled beta distribution defined on $[-1, 1]$ with parameters $[\frac{1}{2}(k+1)]$ and $[\frac{1}{2}k]+1$.

STEP 2 : Replace $\underline{\phi}_p$ in $g(\underline{\phi}_p, \underline{\theta}_q)$ by $\underline{\gamma}_p$ according to the relation (3.2) between $\underline{\phi}_p$ and $\underline{\gamma}_p$.

STEP 3 : Generate γ_k's ($k = 1, 2, \ldots, q$) independently from a rescaled beta distribution defined on $[-1, 1]$ with parameters $[\frac{1}{2}(k+1)]$ and $[\frac{1}{2}k]+1$.

STEP 4 : Replace $\underline{\theta}_q$ in $g(\underline{\phi}_p, \underline{\theta}_q)$ by $\underline{\gamma}_q$ according to the relation (3.2) between $\underline{\theta}_q$ and $\underline{\gamma}_q$.

STEP 5 : After iterating N times from STEP 1 to STEP 4,

$$\frac{1}{N}\sum_{j=1}^{N} g(\underline{\gamma}_{pj}, \underline{\gamma}_{qj})$$

is obtained as an estimate of the integral in (3.1).

4. APPLICATION

We apply the ARMA model identification procedure discussed in this paper to four time series data in Box and Jenkins(1978)' and Wei(1991)'s texts. Assuming the uniform prior for $(\underline{\phi}_p, \underline{\theta}_q)$ in its stationarity-invertibility region and the Jefferys prior and the reference prior for $(\mu, \sigma_\varepsilon)$, the Bayes factors for 15 models, $ARMA(p,q)$, for $(p,q) \in J = \{(i,j) : i = 0,1,2,3,\ j = 0,1,2,3,\ (i,j) \neq (0,0)\}$, are computed from (2.4), respectively, and then the posterior probabilities for each model are directly obtained from (2.5).

For the prior probabilities there are often two alternatives. One is to assign equal probabilities to all models, thus

$$P(ARMA(p,q) \mid \underline{Z}) = \{ \sum_{(p',q')} B_{(p',q')(p,q)} \}^{-1} .$$

Another is to assign unequal probabilities according to the principle of parameters parsimony. Then the prior probability for the $ARMA(p,q)$ model can be assigned as

$$p_{(p,q)} = \frac{(p+q)^{-1}}{\sum_{(i,j)\in J}(i+j)^{-1}} .$$

Table 4.1-4.4 show posterior probabilities computed for 15 ARMA models under the equal and the unequal prior probability for each model. We can know that two priors give about the same results. All the computations are carried using the IMSL subroutines on a UNIX workstation. For the comparisons in each table are shown Akaike(1974)'s AIC criterion and SBC criterion computed through PROC ARIMA procedure of SAS/ETS software. In each table, * is marked for the ARMA model that the maximum likelihood estimation may not converge and ** is marked for ARMA model that the maximum likelihood estimation algorithm did not converge. For such models we often doubt all the results of estimation including AIC and SBC.

The posterior probability of each model is computed and checked every 100 iterations for its stability. Table 4.5 shows the sample size, the number of iterations for the stability, the computer time required and the selected model that gives the maximum posterior probability. Model identification method introduced in this paper requires much more computer time than the AIC criterion or the SBC. But this method can be applied to time series data with the small sample size or the moderate sample size, while results by the maximum likelihood estimation require large sample size since the maximum likelihood estimation for the ARMA model holds under the asymptotic assumption. Also our experiment for a number of simulated data shows that this method select simpler model.

Table 4.1: Posterior probabilities for the ARMA (p, q) model.(Data source: Box and Jenkins(1978)'s Series-E)

ARMA			Equal Probability			Unequal Probability	
(p, q)	AIC	SBC	Jeffreys	Reference	Prior	Jeffreys	Reference
(1,0)	901.81	907.02	0.000	0.000	0.163	0.000	0.000
(2,0)	835.23	843.05	0.073	0.075	0.081	0.112	0.115
(3,0)	832.99	843.41	0.076	0.077	0.054	0.078	0.079
(0,1)	904.97	910.18	0.000	0.000	0.163	0.000	0.000
(0,2)	850.83	858.64	0.000	0.000	0.081	0.000	0.000
(0,3)	841.47	851.89	0.005	0.005	0.054	0.005	0.005
(1,1)	846.82	854.63	0.000	0.000	0.081	0.000	0.000
(1,2)	836.93	847.35	0.051	0.053	0.054	0.052	0.054
(1,3)	837.76	850.79	0.012	0.012	0.041	0.009	0.010
(2,1)	*831.05*	*841.47*	*0.586*	*0.579*	*0.054*	*0.602*	*0.594*
(2,2)	832.68	845.70	0.005	0.005	0.041	0.004	0.004
(2,3)	834.67	850.30	0.029	0.030	0.033	0.018	0.019
(3,1)	834.89	847.92	0.122	0.123	0.041	0.094	0.094
(3,2)	834.09	849.72	0.041	0.041	0.033	0.025	0.025
(3,3)*	836.41	854.64	0.000	0.000	0.027	0.000	0.000

Table 4.2: Posterior probabilities for the ARMA (p, q) model.(Data source: Box and Jenkins(1978)'s Series-F)

ARMA			Equal Probability			Unequal Probability	
(p, q)	AIC	SBC	Jeffreys	Reference	Prior	Jeffreys	Reference
(1,0)	535.96	540.45	0.250	0.252	0.163	0.440	0.441
(2,0)	535.66	542.40	0.128	0.128	0.081	0.113	0.112
(3,0)	537.65	546.64	0.014	0.015	0.054	0.008	0.009
(0,1)	540.03	544.52	0.024	0.025	0.163	0.042	0.043
(0,2)	536.71	543.45	0.173	0.172	0.081	0.152	0.151
(0,3)	537.78	546.77	0.026	0.026	0.054	0.015	0.015
(1,1)	536.14	542.88	0.105	0.105	0.081	0.092	0.092
(1,2)	537.55	546.55	0.043	0.043	0.054	0.025	0.025
(1,3)**			0.044	0.045	0.041	0.019	0.020
(2,1)	537.65	546.65	0.084	0.084	0.054	0.049	0.049
(2,2)	539.54	550.78	0.061	0.061	0.041	0.027	0.027
(2,3)*	541.37	554.86	0.004	0.004	0.033	0.001	0.001
(3,1)	539.44	550.68	0.017	0.017	0.041	0.007	0.007
(3,2)	541.34	554.84	0.026	0.026	0.033	0.009	0.009
(3,3)*	543.33	559.07	0.000	0.000	0.027	0.000	0.000

Table 4.3: Posterior probabilities for the ARMA (p, q) model. (Data source: Wei(1990)'s Series-W1)

ARMA			Equal Probability			Unequal Probability	
(p, q)	AIC	SBC	Jeffreys	Reference	Prior	Jeffreys	Reference
(1,0)	62.07	65.68	0.369	0.367	0.163	0.545	0.542
(2,0)	63.66	69.08	0.121	0.120	0.081	0.089	0.089
(3,0)	65.66	72.89	0.018	0.018	0.054	0.009	0.009
(0,1)	64.46	68.08	0.078	0.079	0.163	0.115	0.117
(0,2)	64.56	69.98	0.040	0.041	0.081	0.030	0.030
(0,3)	66.01	73.23	0.017	0.017	0.054	0.009	0.009
(1,1)	63.69	69.11	0.160	0.160	0.081	0.118	0.118
(1,2)	65.65	72.87	0.064	0.064	0.054	0.031	0.032
(1,3)	67.55	76.59	0.009	0.009	0.041	0.003	0.003
(2,1)	65.66	72.89	0.057	0.056	0.054	0.028	0.028
(2,2)*	66.06	75.09	0.010	0.010	0.041	0.004	0.004
(2,3)*	70.17	81.01	0.003	0.003	0.033	0.001	0.001
(3,1)*	66.34	75.38	0.029	0.030	0.041	0.011	0.011
(3,2)*	68.26	79.10	0.026	0.026	0.033	0.008	0.008
(3,3)	71.07	83.72	0.001	0.001	0.027	0.000	0.000

Table 4.4: Posterior probabilities for the ARMA (p,q) model. (Data source: Wei(1990)'s Series-W5)

ARMA			Equal Probability			Unequal Probability	
(p,q)	AIC	SBC	Jeffreys	Reference	Prior	Jeffreys	Reference
(1,0)	137.89	141.00	0.217	0.221	0.163	0.467	0.473
(2,0)*	139.72	144.38	0.055	0.057	0.081	0.059	0.061
(3,0)	137.76	143.98	0.152	0.150	0.054	0.109	0.107
(0,1)*	191.16	194.27	0.000	0.000	0.163	0.000	0.000
(0,2)	172.30	176.97	0.000	0.000	0.081	0.000	0.000
(0,3)*	158.11	164.33	0.000	0.000	0.054	0.000	0.000
(1,1)	139.79	144.45	0.040	0.041	0.081	0.043	0.044
(1,2)	137.86	144.08	0.179	0.176	0.054	0.128	0.126
(1,3)	139.80	147.58	0.060	0.058	0.041	0.032	0.031
(2,1)	139.89	146.11	0.037	0.038	0.054	0.027	0.027
(2,2)	139.80	147.57	0.101	0.100	0.041	0.055	0.053
(2,3)*	141.81	151.14	0.027	0.027	0.033	0.012	0.012
(3,1)	139.75	147.53	0.113	0.111	0.041	0.061	0.059
(3,2)**			0.012	0.013	0.033	0.005	0.005
(3,3)*	142.76	153.64	0.007	0.008	0.027	0.003	0.003

Table 4.5: Result of ARMA(p,q) model identification.

Time Series Data	Sample Size	Number of iterations	Computer Time	Selected Model
Series-E	100	400	80 minutes	ARMA(2,1)
Series-F	70	100	9 minutes	AR(1)
Series-W1	45	200	5 minutes	AR(1)
Series-W5	35	1000	15 minutes	AR(1)

References

Akaike, H.(1974). "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control,* AC -19, 716-723.

Berger, J. O. and Pericchi, L. R.(1994). "The Intrinsic Bayes Factor for Model

Selection and Prediction," *Technical Report,* 93-43C, Department of Statistics, Purdue University.

Berger, J. O. and Pericchi, L. R.(1996). "The Intrinsic Bayes Factor for Model Selection and Prediction," *Journal of the American Statistical Association* Vol.91, No.433, 109-122.

Box, G.E.P. and Jenkins,G.M.(1976). *Time Series Analysis: forecasting and control*, Holden Day.

Choi,B.S.(1992). *ARMA Model Identification*, Springer-VERLAG, New York.

Jefferys, H.(1961). *Theory of Probability*, London: Oxford University Press.

Jones, M.C.(1987). "Randomly Choosing Parameters from the Stationarity and Invertibility Region of Autoregressive-Moving Average Models," *Applied Statistics,* Vol.36, No.2 134-138.

Leeuw, J.V.D.(1994). "The Covariance Matrix of ARMA Errors in Closed Form," *Journal of Econometrics,* Vol.63, 397-405.

Monahan, J.F.(1983). "Fully Bayesian Analysis of ARMA Time Series Models," *Journal of Econometrics,* Vol.21 307-331.

Monahan, J.F.(1984). "A Note on Enforcing Stationarity in Autoregressive-Moving Average Models," *Biometrika,* Vol.71, No.2, 403-404.

Schwarz, G.(1978) "Estimating the Dimension of a Model," *The Annals of Statistics,* Vol.6, 461-464.

Varshavsky,J.A.(1995). "Intrinsic Bayes Factor for Model Selection with Autoregressive Data," *Bayesian Statistics,* Vol.6.

Wei, W.W.S.(1990). *Time Series analysis*, Addison Wesley.