A BAYESIAN ANALYSIS OF MOVING AVERAGE PROCESSES

By

SAMIR MOUSTAFA SHAARAWY

Bachelor of Science
Cairo University
Cairo, Egypt
1973

Master of Science
Oklahoma State University
Stillwater, Oklahoma
1981

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
July, 1984

A BAYESIAN ANALYSIS OF MOVING AVERAGE PROCESSES

Thesis Approved:

Lyle Broemeling
Thesis Adviser

H. G. Burchard

P. Larry Claypool

Dayll E. Ray

Norman N. Durham
Dean of the Graduate College

ACKNOWLEDGMENTS

I would like to express my appreciation to my major adviser, Professor Lyle D. Broemeling, for his assistance with this problem and for his encouragement throughout my graduate studies. Appreciation is also expressed to other committee members, Dr. Hermann G. Burchard for his suggestions in writing the computer programs and the mathematical proofs, Dr. David L. Weeks for his encouragement and aid, Dr. Daryll E. Ray for his valuable assistance and comments, and Dr. P. Larry Claypool for his help.

Also, I am grateful to Dr. John P. Chandler for his valuable advice and discussions in performing the computer work. I also wish to thank Ms. Barbara Newport for excellent typing. Sincere gratitude is extended to my wife for her patience and sacrifices, and my parents for their interests, encouragement, and support.

Finally, I would like to acknowledge the Office of Naval Research through contract number N00014-82-K-0292.

TABLE OF CONTENTS

Chapter | Page

LIST OF TABLES

Table		Page

CHAPTER I

INTRODUCTION

Although the analysis of time series has recently received great attention, little work has appeared in regard to Bayesian approach to the analysis of time series. Two common ways are known in the statistical literature to analyze time series data. One way concentrates on the analysis of the spectrum of the generating process. Another way is based on the autoregressive moving average (ARMA) parameterization. In general the analysis of ARMA models is a special case of general linear models in which the response or the dependent variable is linearly regressed on a finite number (p) of the previous values of the process and a finite number (q) of the previous random shocks. This means that in ARMA(p,q) processes we have two different groups of parameters; the first group consists of the autoregressive parameters, while the second group consists of the moving average parameters. If the second group contains no elements, we will have a pure autoregressive model of order p, which is denoted by AR(p). If the first group contains no elements, we will have a pure moving average model of order q, which is denoted by MA(q).

With the exception of AR(p) models, very little work has been done from a Bayesian viewpoint in order to analyze ARMA(p,q) processes. The difficulty with moving average processes is that the joint and marginal posterior distributions of the parameters are not standard, thus inferences about the parameters need to be done numerically. A simple anallytic form for the likelihood function is needed, and this has not been done because there is not a closed form for the precision matrix nor for the determinant of the covariance matrix.

The main purpose of this research study is to develop the Bayesian analysis of moving average processes. In the moving average process, the response is linearly regressed on a finite number of independent random "shocks" which have the same probability distribution with zero mean and a constant variance.

In Chapter III, the exact theoretical and analytic forms of the posterior distribution of the MA(1) parameter will be derived, the exact conditional distribution of the error precision will be found, and the exact conditional predictive density of one step ahead forecast will be introduced.

An approximate procedure will be developed to analyze the general moving average processes in Chapter IV. The approximate procedure assumes that the initial values of the "white noise" are constants and equal to their unconditional expectations, namely zero. Although the approximate procedure reduces the number of calculations needed to compute the likelihood function, the Bayesian inferences about the parameters should be done using one of the numerical integration techniques which requires a certain level of mathematical background and much computing time. To overcome this problem, an approximate t distribution will be developed to analyze the general moving average process without using any numerical integration techniques. This will save money and time.

In Chapter V, some numerical problems will be studied to demonstrate the idea of using the exact conditional likelihood function and t approximation procedures. The classical solution of these numerical problems, represented by Box and Jenkins' procedure, will also be given. Chapter VI is a brief summary of the achieved results.

CHAPTER II

REVIEW OF THE LITERATURE

For well-understood reasons, most of the publications devoted to the analysis of ARMA processes using the Bayesian approach concentrate on the analysis of autoregressive processes and pay little attention to moving average processes or to the mixed processes. This void in the Bayesian literature for ARMA models is due to the complexity of the likelihood function of the moving average processes and the mixed processes because there is not a closed form for the likelihood function in terms of the parameters directly.

Wise (1955) and Siddiqui (1958) have obtained the precision matrix for stationary autoregressive processes in closed form. On the other hand, it has not been found, so far, possible to express the precision matrix for the moving average processes in an analytic form in terms of the parameters. Prior to 1970 there were few practical publications on Bayesian approach to analyze time series. Aoki's (1967) book is considered as a Bayesian theoretic study of linear dynamic systems of engineering control, and Zellner (1971) introduces the reader to the Bayesian analysis of autoregressive models. Although the methodology of Box and Jenkins (1970, chapter 7) for identification, estimation, and forecasting is classical, they present Bayesian estimation based on Jeffreys' (1961, section 3.10) prior. They derive an approximate posterior distribution for the MA(q) parameters using their "backforecasting" procedure to attack the initial values problem. They also give an exact expression for the likelihood function of MA(q) processes, and their approach has been independently extended to ARMA by Dent (1977) and Ali (1977). However, the approximation given by Box and Jenkins is restricted by the invertibility conditions and it is difficult to use if $q > 1$.

The forecasting problem from a Bayesian viewpoint of time series data has been discussed by Harrison and Stevens (1971) and they have studied the changes in slope and trends over a short term. Zellner (1971, chapter 7) constructs the Bayesian estimation for the autoregressive and distributed laged models. Newbold (1973) is concerned with Bayesian estimation of the coefficients of the transfer-noise models proposed by Box and Jenkins (1970). He uses a nonlinear least square approximation to show that the Bayesian inferences about the parameters can be done using student's t distribution if Jeffreys' prior is used. In 1974 he derives an exact form for the likelihood function of ARMA processes. The paper written by Harrison and Stevens (1976) is considered as a general review for the time series models which can be analyzed by a Bayesian approach. They found a way by which the time series can be broken into two stages and then use a Bayesian approach. At the second stage, they introduce a random variable θ_t to represent the "level" of the processes. Assuming a normal distribution for θ_0, the posterior distribution of θ_t can be found. A paper by Smith (1979) continued the work of Harrison and Stevens by redefining their steady state model across a large class of sampling distributions that are not normal. McLeod (1977) has proposed replacing the determinant of the coveriance matrix of ARMA processes by its asymptotic limit in order to develop an approximation for the likelihood function. However, his approach does not avoid the problem of computing the precision matrix.

Phadke and Kedem (1978) show three different techniques to obtain the exact likelihood function for MA(q) processes; however, none of these techniques avoid the problem of computing the precision matrix. Their work is extended to ARMA processes by Ansely (1979).

Another approach to approximate the likelihood function of MA(q) processes is used by Zellner and Reynolds (1978) and Hilmer and Tiao (1979). The idea of setting the initial values of the residuals to zero has been used in their work. Zellner and Reynolds show that statistical inferences about the parameters can be approximately done using t distribution by replacing the exact parameters values in the covariance matrix by initial consistent extimates. However, the idea of setting the initial values of the residuals to zero has been used before by Box and Jenkins (1970) and Wilson (1973).

By introducing a new spectral parameterization of time series data, Shore (1980) shows that Whittle's approximation of the likelihood function (1951, chapter 4) can be used in Bayesian analysis of ARMA models. They derive a conjugate prior distribution for their approximation. They show also that the approximate precision matrix of MA(q) is the covariance matrix of AR(p). The paper by Peterka (1981) shows that on a Bayesian basis it is possible to build a consistent theory of system identification. Some other simplifications for the likelihood function are given by Ljung and Box (1976), Nicholls and Hall (1979).

CHAPTER III

THE EXACT BAYESIAN ANALYSIS OF MA(1) PROCESS

The principal goal of this chapter is to develop the exact Bayesian analysis of MA(1) model. An exact analytic form for the posterior distribution of the process coefficient ϕ will be introduced, the exact conditional distribution of the error's precision τ given ϕ will be derived, and the exact conditional predictive density of the next observation Y_{n+1} given ϕ will be constructed. Also, the method of finding the marginal expectations and variances of τ^{-1} and Y_{n+1} will be discussed.

3.1 Definitions and Notations

Let $\{t\}$ be a sequence of integers, ϕ be a real constant, $\{\varepsilon_t\}$ be a sequence of "white noise", and Y_t be the realization of the process $\{Y_t\}$ at time t. Then the moving avarage model of the first order is defined by

$$Y_t = \varepsilon_t - \phi\varepsilon_{t-1}, \qquad t = \ldots,-1,0,1,\ldots.$$

Assuming that we have n observations, the MA(1) model can be written in matrix notation as

$$\underset{\sim}{Y} = A(\phi)\underset{\sim}{\varepsilon}$$

where

$$\underset{\sim}{Y} = (Y_1 , Y_2 , Y_3 , \dots , Y_n)^T,$$

$$\underset{\sim}{\varepsilon} = (\varepsilon_0 , \varepsilon_1 , \varepsilon_2 , \dots , \varepsilon_n)^T,$$

and

$$A(\phi) = \begin{bmatrix} -\phi & 1 & & & & \Phi & \\ & -\phi & 1 & & & & \\ & & \cdot & \cdot & & & \\ & & & \cdot & \cdot & & \\ & \Phi & & & \cdot & \cdot & \\ & & & & & -\phi & 1 \end{bmatrix}_{n \times (n+1)} .$$

Assume that the errors are independent and normally distributed with zero means and a constant variance τ^{-1}. Thus,

$$\underset{\sim}{Y} \sim \text{Normal}\ (\underset{\sim}{0},\ (A(\phi)\text{Var}(\underset{\sim}{\varepsilon})A'(\phi))^{-1})$$

i.e.,

$$\underset{\sim}{Y} \sim N(\underset{\sim}{0},\ (\tau^{-1}A(\phi)A'(\phi))^{-1}). \tag{3.1.1}$$

From (3.1.1) we can write the density of $\underset{\sim}{Y}$ given the parameters (ϕ,τ) as

$$f(\underset{\sim}{Y}|\phi,\tau) = (2\pi)^{-\frac{n}{2}}|\tau[A(\phi)A'(\phi)]^{-1}|^{\frac{1}{2}}\ \text{Exp}\{-\tfrac{\tau}{2}\underset{\sim}{Y}'[A(\phi)A'(\phi)]^{-1}\underset{\sim}{Y}\}, \quad \underset{\sim}{Y} \in R^n,$$

$$\phi \in R, \quad \tau > 0,$$

$$= 0, \quad \text{otherwise}, \tag{3.1.2}$$

i.e., the likelihood function of (ϕ,τ) is

$$L(\phi,\tau|\underset{\sim}{Y}) \propto \tau^{\frac{n}{2}}|A(\phi)A'(\phi)|^{-\frac{1}{2}}\ \text{Exp}\{-\tfrac{\tau}{2}\underset{\sim}{Y}'[A(\phi)A'(\phi)]^{-1}\underset{\sim}{Y}\}, \quad \phi \in R, \quad \tau > 0,$$

$$Y \in R^n,$$

$= 0, \quad$ otherwise. (3.1.3)

3.2 An Algorithm to Develop the Exact Analysis of MA(1) Process

We have mentioned before that, regardless of the form of the prior density, the posterior distributions are not standard, thus statistical inferences about the parameters should be done numerically. However, the form (3.1.3) of the likelihood function is useless in developing a practical Bayesian analysis of MA(1) process. To evaluate $L(\phi,\tau|\underset{\sim}{Y})$ for a specific value of ϕ, we have to compute $|A(\phi)A'(\phi)|$ and $(A(\phi)A'(\phi))^{-1}$. To overcome these two problems in developing the Bayesian analysis of MA(1) model, an orthogonal matrix Q will be constructed such that $Q'A(\phi)\cdot A'(\phi)Q = D$ is a diagonal matrix. The diagonal elements of D are the eigenroots of the matrix $A(\phi)A'(\phi)$; the eigenroots will be given in closed form. It will be shown that the elements of Q are independent of the parameter ϕ which saves much time in computing $L(\phi,\tau|\underset{\sim}{Y})$.

By constructing the matrix Q and the eigenroots λ_i's of $A(\phi)A'(\phi)$, it will be possible to develop a theoretical and practical Bayesian analysis of MA(1) process.

3.2.1 The Analytic Forms of Q and λ_i's

The matrix $A(\phi)A'(\phi)$ is symmetric positive definite, thus their exists an orthogonal matrix Q such that

$$Q'A(\phi)A'(\phi)Q = \text{Diag}(\lambda_1(\phi),\lambda_2(\phi),\ldots,\lambda_n(\phi))$$

where λ_i's are the eigenroots of the matrix $V(\phi) = A(\phi)A'(\phi)$ which is a tridiagonal Toeplitz matrix with the following form:

$$V(\phi) = \begin{bmatrix} 1+\phi^2 & -\phi & & & \Phi \\ -\phi & \cdot & \cdot & \cdot & \\ & \cdot & \cdot & \cdot & \\ & \cdot & \cdot & \cdot & \\ \Phi & & & & -\phi \\ & & & -\phi & 1+\phi^2 \end{bmatrix}_{n\times n} .$$

It can be shown that (Grenander and Szegö, 1958)

$$\lambda_t(\phi) = 1 + \phi^2 - 2\phi\cos(\frac{t\pi}{n+1}), \qquad t = 1,2,\ldots,n.$$

Let $\underset{\sim}{x}_t = (x_1,x_2,\ldots,x_n)^T$ be the eigenvector corresponding to λ_t, then

$$A(\phi)A'(\phi)\underset{\sim}{x}_t = \lambda_t\underset{\sim}{x}_t, \qquad t = 1,2,\ldots,n$$

i.e.

$$V(\phi)\underset{\sim}{x}_t = \lambda_t\underset{\sim}{x}_t, \qquad t = 1,2,\ldots,n. \tag{3.2.1.1}$$

Thus, for every t we have the following system of equations:

$$x_2 = 2x_1\cos(\frac{t\pi}{n+1}) \tag{1}$$

$$x_3 = 2x_2\cos(\frac{t\pi}{n+1}) - x_1 \tag{2}$$

$$x_4 = 2x_3\cos(\frac{t\pi}{n+1}) - x_2 \tag{3}$$

$$\vdots \qquad\qquad\qquad\qquad \vdots$$

$$x_{r+2} = 2x_{r+1}\cos(\frac{t\pi}{n+1}) - x_r \tag{r+1}$$

$$\vdots \qquad\qquad\qquad\qquad \vdots$$

$$x_n = 2x_{n-1}\cos(\frac{t\pi}{n+1}) - x_{n-2} \tag{n-1}$$

$$x_{n-1} = 2x_n\cos(\frac{t\pi}{n+1}) \tag{n}$$

The above system is a linear system in n unknowns and (n-1) independent equations. The system can be written as

$$A^* \underset{\sim}{x} = \underset{\sim}{0}$$

where $A^* = A(\phi)A'(\phi) - \lambda_t I$.

Claim: The matrix A^* has rank (n-1).

Proof: $|A^*| = 0 \Longrightarrow$ rank $(A^*) = R(A^*) \leq n-1$. Let $\Delta_n = |A^*|$. It is easy to see that the following difference equation is satisfied:

$$\Delta_{n-1} = s\Delta_{n-2} - \Delta_{n-3}\ , \qquad s = 2\cos(\frac{t\pi}{n+1}). \qquad (3.2.1.2)$$

It can be shown that the solution of the difference equation is

$$\Delta_{n-1} = \sin(\frac{nt\pi}{n+1})/\sin(\frac{t\pi}{n+1}), \qquad t = 1,2,\ldots,n$$

i.e.,

$$\Delta_{n-1} \neq 0 \Longrightarrow R(A^*) = n-1. \qquad \text{Q.E.D.}$$

So, the system of equations (3.2.1.1) has one independent solution.

Let

$$x_1 = \sqrt{\frac{2}{n+1}}\ \sin(\frac{t\pi}{n+1}).$$

It can be shown that

$$x_{r+2} = \sqrt{\frac{2}{n+1}}\ \sin(\frac{(r+2)t\pi}{n+1}), \qquad r = 1,2,\ldots,n-2.$$

Thus, the eigenvector corresponding to λ_t is

$$\underset{\sim}{x}_t = \sqrt{\frac{2}{n+1}}\ (\sin(\frac{t\pi}{n+1})\ \sin(\frac{2t\pi}{n+1})\ \ldots\ \sin(\frac{nt\pi}{n+1}))^T, \quad t=1,2,\ldots,n \qquad (3.2.1.3)$$

i.e.,

$$(q_{ij}) = Q = \sqrt{\frac{2}{n+1}}\,\sin\left(\frac{ij\pi}{n+1}\right), \qquad i,j = 1,2,\ldots,n. \tag{3.2.1.4}$$

3.2.2 The Exact Joint Distribution of (ϕ,τ)

Let the prior density of ϕ and τ be given by

$$\xi(\phi,\tau) \propto \tau^{a-1}\, e^{-\tau b}\, \xi_1(\phi), \qquad \tau > 0,\ \phi \in R,\ a > 0,\ b > 0.$$

Where $\xi_1(\phi)$ is the marginal prior density of ϕ, and $\xi_1(\phi)$ can be a proper density. The hyperparameter domain of (a,b) can be extended to include the point (0,0) if one uses Jeffreys' prior for τ. Regardless of the choice of $\xi_1(\phi)$, the posterior distribution of (ϕ,τ) can be written as

$$P(\phi,\tau) \propto \tau^{\frac{n+2a}{2}-1}\, \xi_1(\phi)\left(\prod_{t=1}^{n} \lambda_t^{-\frac{1}{2}}(\phi)\right) \mathrm{Exp}\{-\frac{\tau}{2}[\sum_{j=1}^{n} w_j^2\lambda_j^{-1}(\phi) + 2b]\}, \quad \phi \in R,$$

$$\tau > 0,\ \underset{\sim}{Y} \in R^n, \tag{3.2.2.1}$$

where $\underset{\sim}{W} = (w_1, w_2, \ldots, w_n)^T = Q\underset{\sim}{Y} \in R^N$.

3.3 The Exact Marginal Distribution of ϕ

$$P_1(\phi|\underset{\sim}{Y}) = \int_0^\infty P(\phi,\tau|\underset{\sim}{Y})\,d\tau$$

Using (3.2.2.1)

$$P_1(\phi|\underset{\sim}{Y}) \propto \int_0^\infty \tau^{\frac{n+2a}{2}-1}\, \xi_1(\phi)\left(\prod_{t=1}^{n} \lambda_t^{-\frac{1}{2}}(\phi)\right) \mathrm{Exp}\{-\frac{\tau}{2}[\sum_{j=1}^{n} w_j^2\lambda_j^{-1}(\phi) + 2b]\}\,d\tau.$$

So,

$$P_1(\phi|\underset{\sim}{Y}) \propto \xi_1(\phi)\left(\prod_{t=1}^{n} \lambda_t^{-\frac{1}{2}}(\phi)\right) \int_0^\infty \tau^{\frac{n+2a}{2}-1}\, \mathrm{Exp}\{-\frac{\tau}{2}[\Sigma w_j^2\lambda_j^{-1}(\phi) + 2b]\}\,d\tau$$

$$\propto \xi_1(\phi)\left(\prod_{t=1}^{n}\lambda_t^{-\frac{1}{2}}(\phi)\right)\Gamma\left(\frac{n+2a}{2}\right)\Big/\left[\frac{\Sigma w_j^2\lambda_j^{-1}(\phi)+2b}{2}\right]^{\frac{n+2a}{2}}$$

Finally,

$$P_1(\phi|\underset{\sim}{Y}) \propto \xi_1(\phi)\prod_{t=1}^{n}\lambda_t^{-\frac{1}{2}}(\phi)/[\Sigma w_j^2\lambda_j^{-1}(\phi)+2b]^{\frac{n+2a}{2}}, \qquad \phi \in R. \tag{3.3.1}$$

Although that form (3.3.1) is not standard form, it can be effectively used in practical situations.

3.4 The Exact Conditional Distribution of $\tau|\phi$

$$P(\phi,\tau|\underset{\sim}{Y}) = P(\tau|\phi,\underset{\sim}{Y})P_1(\phi|\underset{\sim}{Y}) \tag{3.4.1}$$

Comparing (3.4.1) with (3.2.2.1), we conclude that

$$P(\tau|\phi,\underset{\sim}{Y}) \propto \tau^{\frac{n+2a}{2}-1}\operatorname{Exp}\{-\frac{\tau}{2}[\Sigma w_j^2\lambda_j^{-1}(\phi) + 2b]\}, \quad \tau>0,\ \phi\in R,\ \underset{\sim}{Y}\in R^n \tag{3.4.2}$$

i.e., the conditional posterior distribution of τ given ϕ is gamma with parameters

$$(n+2a)/2 \quad \text{and} \quad (\Sigma w_j^2\lambda_j^{-1}(\phi)+2b)/2.$$

Let the expectation and the variance of a random variable X be denoted by

$E(X)$ and $V(X)$, respectively.

Although that it is difficult to find the marginal distribution of $\tau|\underset{\sim}{Y}$, the marginal expectation and variance of $\tau^{-1}|\underset{\sim}{Y}$ can be numerically obtained from the following relations:

$$\underset{\tau}{E}(\tau^{-1}|\underset{\sim}{Y}) = \underset{\phi}{E}\ \underset{\tau}{E}\ (\tau^{-1}|\phi,\underset{\sim}{Y}) = \int_R\left(\frac{\Sigma w_j^2\lambda_j^{-1}(\phi)+2b}{n+2a-2}\right)P_1(\phi|\underset{\sim}{Y})d\phi \tag{3.4.3}$$

$$\underset{\tau}{V}(\tau^{-1}|\underset{\sim}{Y}) = \underset{\phi}{E}\ \underset{\tau}{V}\ (\tau^{-1}|\phi,\underset{\sim}{Y}) + \underset{\phi}{V}\ \underset{\tau}{E}\ (\tau^{-1}|\phi,\underset{\sim}{Y}) \tag{3.4.4}$$

where

$$\underset{\tau}{V}\,(\tau^{-1}|\phi,\underset{\sim}{Y}) = \frac{2}{n+2a-4}[E(\tau^{-1}|\phi,\underset{\sim}{Y})]^2$$

and

$$\underset{\phi}{V}\,\underset{\tau}{E}\,(\tau^{-1}|\phi,\underset{\sim}{Y}) = \underset{\phi}{E}\,[E(\tau^{-1}|\phi,\underset{\sim}{Y})]^2 - [E(\tau^{-1}|\underset{\sim}{Y})]^2.$$

3.5 The Exact Conditional Distribution of $Y_{n+1}|\phi$

Often, the final goal of time series is forecasting. One part of the forecasting problem is to make inferences about the unobserved observations given the history of the process $\{Y_\tau\}$. The predictive density of the unobserved observations is the Bayesian tool to solve this problem. However, the predictive density can be viewed as an average of the conditional predictive densities. In this section, it will be shown that it is possible to find an exact form for the conditional predictive density of the next observation Y_{n+1}, while the marginal predictive density should be numerically done.

Let $P(Y_{n+1}|\underset{\sim}{Y})$ be the predictive density of Y_{n+1}, then

$$\begin{aligned}
P(Y_{n+1}|\underset{\sim}{Y}) &= \int_\phi \int_\tau P(Y_{n+1},\phi,\tau|\underset{\sim}{Y})\,d\tau d\phi \\
&= \int_\phi \int_\tau P(\phi,\tau|\underset{\sim}{Y})P(Y_{n+1}|\phi,\tau,\underset{\sim}{Y})\,d\tau d\phi \\
&\propto \int_\phi \int_\tau P(Y_1,Y_2,\ldots,Y_n|\phi,\tau)\xi(\phi,\tau)P(Y_{n+1}|\phi,\tau)\,d\tau d\phi \\
&\propto \int_\phi \int_\tau P(Y_1,Y_2,\ldots,Y_n,Y_{n+1}|\phi,\tau)\xi(\phi,\tau)\,d\tau d\phi \\
&\propto \int_\phi \int_\tau \tau^{\frac{n+1+2a}{2}-1}\,\mathrm{Exp}-\frac{\tau}{2}[\underset{\sim}{Y}^{*\prime}B(\phi)\underset{\sim}{Y}^{*}+2b]\xi_1(\phi)\prod_{j=1}^{n+1}\lambda_j^{*\,-\frac{1}{2}}(\phi)\,d\tau d\phi
\end{aligned}$$

where $\underset{\sim}{Y}^* = [Y_1, Y_2, \ldots, Y_{n+1}]^T$; $B(\phi) = Q^* \text{Diag}\,(\lambda_j^{*-1}(\phi))Q^*$ and

$$\lambda_j^*(\phi) = 1 + \phi^2 - 2\phi\cos(\frac{j\pi}{n+2}).$$

However, $Q^* = (q_{ij}^*) = \sqrt{\frac{2}{n+2}}\,\sin(\frac{ij\pi}{n+2})$, $i,j = 1,2,\ldots,n+1$. $\underset{\sim}{Y}^{*\prime}B(\phi)\underset{\sim}{Y}$ can be written as $Y_{n+1}^2 d(\phi) + 2Y_{n+1}e(\phi) + c(\phi)$ where

$$d(\phi) = b_{n+1\ n+1}(\phi)$$

$$e(\phi) = \sum_{i=1}^{n} Y_i b_{i\ n+1}(\phi) \quad \text{and}$$

$$c(\phi) = \sum_{i=1}^{n} Y_i^2 b_{ii}(\phi) + 2\sum_{i<j}^{n}\sum Y_i Y_j b_{ij}(\phi) + 2b.$$

Thus,

$$P(Y_{n+1}|\phi,\underset{\sim}{Y}) \propto \xi_1(\phi) \prod_{j=1}^{n+1} \lambda_j^{*\,-\frac{1}{2}}(\phi)/[(Y_{n+1}-d^{-1}(\phi)e(\phi))^2 d(\phi)+c(\phi) -e^2(\phi)d^{-1}(\phi)]^{\frac{n+1+2a}{2}} \tag{3.5.1}$$

i.e., the conditional posterior distribution of Y_{n+1} given ϕ is t with $(n+2a)$ degrees of freedom, $d^{-1}(\phi)e(\phi)$ location parameter, and $\frac{d(\phi)(n+2a)}{c(\phi)-e^2(\phi)d^{-1}(\phi)}$ precision. However, the marginal expectation and variance of $Y_{n+1}|\underset{\sim}{Y}$ can be numerically obtained from the following relations:

$$E(Y_{n+1}|\underset{\sim}{Y}) = \underset{\phi}{E}\,(d^{-1}(\phi)e(\phi)) \tag{3.5.2}$$

$$V(Y_{n+1}|\underset{\sim}{Y}) = \underset{\phi}{E}\,V\,(Y_{n+1}|\phi,\underset{\sim}{Y}) + \underset{\phi}{V}\,E\,(Y_{n+1}|\phi,\underset{\sim}{Y}) \tag{3.5.3}$$

where

$$V(Y_{n+1}|\phi,\underset{\sim}{Y}) = \frac{c(\phi)-e^2(\phi)d^{-1}(\phi)}{d(\phi)(n+2a-2)}\ , \text{ and}$$

$$\underset{\phi}{V}\, E\,(Y_{n+1}|\phi,\underset{\sim}{Y}) = \underset{\phi}{E}\,[E(Y_{n+1}|\phi,\underset{\sim}{Y})]^2 - [E\,(Y_{n+1}|\underset{\sim}{Y})]^2.$$

3.6 Conclusion

The main objective of this chapter was to develop an exact practical analysis of MA(1) process using the Bayesian approach. It has been shown that the marginal posterior distribution of ϕ is not standard but it can be written in an analytic form, using the matrix Q and the eigenroots λ_i's, which can be effectively used in practical situations. Also it is shown that the exact conditional distribution of τ given ϕ is gamma with parameters

$$\frac{n+2a}{2} \quad \text{and} \quad \frac{\Sigma w_j^2 \lambda_j^{-1}(\phi) + 2b}{2}$$

and that the exact conditional distribution of Y_{n+1} given ϕ is a general univariate t with parameters

$$n+2a\ , \quad d^{-1}(\phi)e(\phi), \quad \text{and} \quad \frac{d(\phi)(n+2a)}{c(\phi)-e^2(\phi)d^{-1}(\phi)}\ .$$

Also, it is shown that the marginal expectations and variances of $\tau^{-1}|\underset{\sim}{Y}$ and $Y_{n+1}|\underset{\sim}{Y}$ can be numerically done.

CHAPTER IV

BAYESIAN ANALYSIS OF MOVING AVERAGE PROCESSES

The main problem in analyzing the MA(q) models is that the likelihood function is analytically interactable because there is no closed form for the precision matrix or for the determinent of the covariance matrix in terms of the parameters. Thus, too many calculations are required in order to compute the likelihood function. Furthermore, with the usual prior distribution, the posterior distributions are not standard which means that inference about the parameters must be done numerically. Thus, as the sample size n increases, computation of the likelihood function becomes increasingly laborious even for high speed computers (Shore, 1980).

A complete Bayesian analysis is not possible without finding a way to represent the likelihood function in such a way to produce analytically tractable posterior distributions. Shore (1980) uses Whittle's (1951, chapter 4) approximation of the likelihood function. Although Whittle's approximation reduces the number of needed calculations to characterize the posterior distributions, it requires the validity of stationarity and invertibility assumptions; furthermore, the theoretical and numerical properties of Whittle's approximation have not been thoroughly studied.

Another approach is to approximate the likelihood function of MA(q) process as was done by Hilmer and Tiao (1979). Their approximation depends on setting the initial values of the errors to zeroes. However, their procedure requires the validity of invertibility conditions; furthermore, their approach is not very effective in computing the residual sum of squares. The same idea has been used by Box and Jenkins (1970) and Wilson (1973). Some other investigations which attempt to simplify the likelihood function can be found in Newbold (1974), Ljung and Box (1970), and Nicholls and Hall (1979).

All the above investigations and others try to approximate the likelihood function of MA(q) models by adding additional assumptions to the usual ones in order to produce an analysis of MA(q) processes.

In this chapter, an approximate Bayesian theory will be developed to analyze MA(q) model by adding q more assumptions like the ones used by Hilmer and Tiao (1979). The approximate procedure assumes that the initial values of the errors are constants and equal to their expectations, namely zeroes. Thus, we assume that

$$\underset{\sim}{\varepsilon}^*_{1\times q} = (\varepsilon_{1-q},\varepsilon_{2-q},\ldots,\varepsilon_0)^T = \underset{\sim}{0}.$$

In most of practical situations, q is not more than 2. If $q=2$, this means that two of the errors $\{\varepsilon_t\}$ are assumed to be zero. This gives an approximate covariance structure for the first two observations (y_1,y_2), while the Whittle's approximation gives an approximate covariance structure for all observations $(y_1,y_2,\ldots,y_n)$. Thus as n increases, the recommended approximation to analyze MA(q) models is expected to be effective and reasonable. An algorithm to build a complete Bayesian theory of MA(q) analysis will be introduced in section 4.1; the posterior distribution of the model parameters will be derived using the conditional likelihood function and a normal-gamma density as a prior distribution. In section 4.2, it will be shown that Bayesian inferences about the parameters can be approximately done using a t distribution.

4.1 The Posterior Analysis of MA(q) Processes

The main objective of this section is to develop a general methodology based on equating the initial values of the errors to zeroes in order to achieve a convenient form for the posterior distribution of the model parameters. This is equivalent to basing the likelihood function on the conditional distribution of $y_{q+1}, y_{q+2}, \ldots, y_n$ given $y_1, y_2, \ldots, y_q$.

Using the same notation as was used in Chapter III, let $\underset{\sim}{\phi} = (\phi_1, \phi_2, \ldots, \phi_q)^T$ be a vector of real constants, then we can write the MA(q) model approximately as

$$y_t = \varepsilon_t - \sum_{j=1}^{q} \phi_j \varepsilon_{t-j}, \qquad \varepsilon_{1-q} = \varepsilon_{2-q} = \ldots = \varepsilon_0 = 0 \tag{4.1.1}$$

where $t = 1,2,\ldots,n$.

In matrix notation, the model (4.1.1) can be written

$$\underset{\sim}{y}_{n\times 1} = A(\underset{\sim}{\phi})\underset{\sim}{\varepsilon}_{n\times 1}$$

where

$$\underset{\sim}{y}_{n\times 1} = (y_1, y_2, \ldots, y_n)^T, \quad \underset{\sim}{\varepsilon}_{n\times 1} = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)^T, \quad \text{and}$$

$$A(\underset{\sim}{\phi}) = \begin{bmatrix} 1 & & & & & \\ -\phi_1 & & & & & \\ -\phi_2 & & & & \Phi & \\ \vdots & & \ddots & & & \\ -\phi_q & & \ddots & \ddots & & \\ \Phi & \ddots & & & & \\ & & -\phi_q & -\phi_2 & -\phi_1 & 1 \end{bmatrix}.$$

However, the covariance matrix of y is $\frac{1}{\tau}A(\underset{\sim}{\phi})A'(\underset{\sim}{\phi})$ where τ is the error precision. Clearly,

$$|A(\phi)A'(\underset{\sim}{\phi})| = 1,$$

thus the conditional density of $y_{q+1}, y_{q+2}, \ldots, y_n$ given $y_1, y_2, \ldots, y_q$, ϕ and τ can be written as

$$L(\underset{\sim}{\phi},\tau) \propto \tau^{\frac{n}{2}} e^{-\frac{\tau}{2}\underset{\sim}{\varepsilon}'(\underset{\sim}{\phi})\underset{\sim}{\varepsilon}(\underset{\sim}{\phi})}, \qquad \underset{\sim}{\phi} \in R^q, \quad \tau > 0, \quad \underset{\sim}{y} \in R^n. \tag{4.1.2}$$

Thus the problem of developing a practical form for the likelihood function becomes the problem of developing a method by which the exponent $\underset{\sim}{\varepsilon}'(\underset{\sim}{\phi})\underset{\sim}{\varepsilon}(\underset{\sim}{\phi})$ can be easily computed.

From (4.1.1) we have

$$\varepsilon_t = y_t + \sum_{j=1}^{q} \phi_j \varepsilon_{t-j}. \tag{4.1.3}$$

Thus (4.1.2) can be written as

$$L(\underset{\sim}{\phi},\tau) \propto \tau^{\frac{n}{2}} \exp -\frac{\tau}{2} \sum_{t=1}^{n} [y_t + \sum_{j=1}^{q} \phi_j \varepsilon_{t-j}]^2, \qquad \underset{\sim}{\phi} \in R^q, \quad \tau > 0. \tag{4.1.4}$$

In (4.1.4) notice the ε_{t-j}'s are defined recursively from (4.1.3), so that for each positive integer s, ε_s is a function of $y_1, y_2, \ldots, y_s$ and $\underset{\sim}{\phi}$ where $1 \leq s \leq n$. To emphasize that ε_{t-j} is a function of $\underset{\sim}{\phi}$, we will write ε_{t-j} as $\varepsilon_{t-j}(\underset{\sim}{\phi})$.

Assume that the prior distribution of the parameters $\underset{\sim}{\phi}$ and τ is a normal gamma in the form

$$\xi(\underset{\sim}{\phi},\tau) \propto \tau^{\frac{q}{2}} e^{-\frac{\tau}{2}(\underset{\sim}{\phi}-\underset{\sim}{\mu})'Q(\underset{\sim}{\phi}-\underset{\sim}{\mu})} \tau^{a-1} e^{-\tau b}, \qquad \underset{\sim}{\phi} \in R^q, \quad \tau > 0, \tag{4.1.5}$$

$$\underset{\sim}{\mu} \in R^q, \quad a > 0, \quad b > 0,$$

where Q is q×q positive definite matrix. However, the domain of the hyperparameters can be extended to include Jeffreys' prior by setting

$a = -\frac{q}{2}$, $b = 0$ and $Q = 0_{q\times q}$. Combining (4.1.4) and (4.1.5), we can write the joint posterior distribution of $\underset{\sim}{\phi}$ and τ as

$$\xi(\underset{\sim}{\phi},\tau|\underset{\sim}{Y}) \propto \tau^{\frac{n+2a+q}{2}-1} \exp -\frac{\tau}{2}[2b + (\underset{\sim}{\phi}-\underset{\sim}{\mu})'Q(\underset{\sim}{\phi}-\underset{\sim}{\mu})$$

$$+ \sum_{t=1}^{n} (y_t + \sum_{j=1}^{q} \phi_j \varepsilon_{t-j}(\underset{\sim}{\phi}))^2], \qquad \underset{\sim}{\phi} \in R^q, \quad \tau > 0 . \tag{4.1.6}$$

Consequently, the marginal posterior distribution of $\underset{\sim}{\phi}$ is

$$\xi(\underset{\sim}{\phi}|\underset{\sim}{Y}) = \int_0^\infty \xi(\underset{\sim}{\phi},\tau|\underset{\sim}{y})d\tau \propto \frac{1}{[2b+(\underset{\sim}{\phi}-\underset{\sim}{\mu})'Q(\underset{\sim}{\phi}-\underset{\sim}{\mu}) + \sum_{t=1}^{n} (y_t + \sum_{j=1}^{q} \phi_j \varepsilon_{t-j}(\underset{\sim}{\phi}))^2]^{\frac{n+2a+q}{2}}},$$

$$\underset{\sim}{\phi} \in R^q. \tag{4.1.7}$$

Although it is easy to compute (4.1.7) recursively, we can not separate the parameters which means that we should use a numerical integration technique to characterize the posterior distribution $\xi(\underset{\sim}{\phi}|\underset{\sim}{Y})$. As q increases, the previous procedure will be expensive and difficult to use; thus the need of having a convenient and analytical approximation for the posterior distributions becomes essential.

4.2 The t Approximation to Analyze MA(q) Models

As we have mentioned, the form (4.1.7) is still complicated and not very useful in making inferences about the parameters. The main reason is that ε_{t-j} is a function of $\underset{\sim}{\phi}$. Our goal here is to develop a t approximation for the marginal distribution of $\underset{\sim}{\phi}$ given by (4.1.7). We notice that if ε_{t-j} is not a function of $\underset{\sim}{\phi}$, the marginal distribution of $\underset{\sim}{\phi}$ would be a t distribution in q dimensions; then our goal is to estimate the residual ε_t, $t = 1,2,\ldots,n$.

An easy way to estimate the residuals is to search in the parameter space for the value $\underset{\sim}{\phi}_0 = (\phi_{10},\phi_{20},\ldots,\phi_{q0})$ which minimizes the residual sum of squares $\underset{\sim}{\varepsilon}'(\underset{\sim}{\phi})\underset{\sim}{\varepsilon}(\underset{\sim}{\phi})$ and then use this value to estimate the residuals. By estimating the residuals, we can write the marginal distribution of $\underset{\sim}{\phi}$ as

$$\xi(\underset{\sim}{\phi}|\underset{\sim}{Y}) \propto \frac{1}{[(\underset{\sim}{\phi}-A^{-1}B)'A(\underset{\sim}{\phi}-A^{-1}B) + C - B'A^{-1}B]^{\frac{n+2a+q}{2}}}, \qquad \underset{\sim}{\phi} \in R^q, \qquad (4.2.1)$$

where

$$A = Q + A^*_{q\times q}$$

$$\underset{\sim}{B} = Q\underset{\sim}{\mu} - \underset{\sim}{B}^*_{q\times 1}$$

$$C = \sum_{t=1}^{n} y_t^2 + \underset{\sim}{\mu}'Q\underset{\sim}{\mu} + 2b$$

and

$$A^*_{q\times q} = (a^*_{ij}) = (\sum_{t=1}^{n} \varepsilon_{t-i}\varepsilon_{t-j}), \qquad i,j = 1,2,\ldots,q$$

$$\underset{\sim}{B}^*_{q\times 1} = (b^*_j) = (\sum_{t=1}^{n} y_t\varepsilon_{t-j}), \qquad j = 1,2,\ldots,q.$$

Thus $\underset{\sim}{\phi}$ has approximately a t distribution in q dimensions with n+2a degrees of freedom, $A^{-1}\underset{\sim}{B}$ location parameter, and $\frac{A(n+2a)}{C-\underset{\sim}{B}'A^{-1}\underset{\sim}{B}}$ precision.

4.3 The Approximate Marginal Distribution of τ

Using the estimated residuals, we can write (4.1.6) as

$$\xi(\underset{\sim}{\phi},\tau|\underset{\sim}{Y}) \propto \tau^{\frac{n+2a+q}{2}-1} \exp -\frac{\tau}{2}[(\underset{\sim}{\phi}-A^{-1}\underset{\sim}{B})'A^{-1}(\underset{\sim}{\phi}-A^{-1}\underset{\sim}{B}) + C - \underset{\sim}{B}'A^{-1}\underset{\sim}{B}],$$

$$\phi \in R^q, \quad \tau > 0.$$

Let the marginal posterior distribution of τ be denoted by $\xi(\tau|\underset{\sim}{Y})$, then

$$\xi(\tau|\underset{\sim}{Y}) = \int_{R^q} \xi(\underset{\sim}{\phi},\tau|\underset{\sim}{Y})d\underset{\sim}{\phi}$$

$$\propto \tau^{\frac{n+2a+q}{2}-1} e^{-\frac{\tau}{2}[C-\underset{\sim}{B}'A^{-1}\underset{\sim}{B}]} \int_{R^q} e^{-\frac{\tau}{2}(\underset{\sim}{\phi}-A^{-1}\underset{\sim}{B})'A^{-1}(\underset{\sim}{\phi}-A^{-1}\underset{\sim}{B})} d\underset{\sim}{\phi}$$

$$\propto \tau^{\frac{n+2a}{2}-1} e^{-\frac{\tau}{2}[C-\underset{\sim}{B}'A^{-1}\underset{\sim}{B}]}, \qquad \tau > 0, \qquad (4.3.1)$$

i.e., the marginal posterior distribution of τ is gamma with parameters

$$\frac{n+2a}{2} \text{ and } \frac{C-\underset{\sim}{B}'A^{-1}\underset{\sim}{B}}{2}.$$

4.4 The Predictive Density of y_{n+1}

Let the joint distribution of $y_1,y_2,\ldots,y_n,y_{n+1}$, $\underset{\sim}{\phi}$, and τ be denoted by $p(y_1,y_2,\ldots,y_n,y_{n+1},\underset{\sim}{\phi},\tau)$, then

$$p(y_1,\ldots,y_{n+1},\underset{\sim}{\phi},\tau) = p(y_1,y_2,\ldots,y_{n+1}|\underset{\sim}{\phi},\tau)\xi(\underset{\sim}{\phi},\tau)$$

$$\propto \tau^{\frac{n+2a+q+1}{2}-1} \exp -\frac{\tau}{2}\{2b + (\underset{\sim}{\phi}-\underset{\sim}{\mu})'Q(\underset{\sim}{\phi}-\underset{\sim}{\mu}) + \sum_{t=1}^{n}[y_t + \Sigma\phi_j\varepsilon_{t-j}(\underset{\sim}{\phi}_0)]^2$$

$$+ [y_{n+1} + \Sigma\phi_j\varepsilon_{n+1-j}(\underset{\sim}{\phi}_0)]^2$$

$$\propto \tau^{\frac{n+2a+q+1}{2}-1} \exp -\frac{\tau}{2}\{2b + (\underset{\sim}{\phi}-\underset{\sim}{\mu})'Q(\underset{\sim}{\phi}-\underset{\sim}{\mu}) + \sum_{t=1}^{n+1}[y_t + \sum_{j=1}^{q}\phi_j\varepsilon_{t-j}(\underset{\sim}{\phi}_0)]^2. \quad (4.4.1)$$

Let

$$H_{q\times q} = (h_{ij}) = \left(\sum_{t=1}^{n+1} \varepsilon_{t-i}(\underset{\sim}{\phi}_0)\varepsilon_{t-j}(\underset{\sim}{\phi}_0)\right), \qquad i,j = 1,2,\ldots,q,$$

$$\underset{\sim}{L}_{q\times 1} = (\ell_i) = (\varepsilon_{n+1-i}(\underset{\sim}{\phi}_0)), \qquad i = 1,2,\ldots,q,$$

$$G_{q\times q} = Q_{q\times q} + H_{q\times q},$$

$$\underset{\sim}{Z}_{q\times 1} = \underset{\sim}{B} - y_{n+1}\underset{\sim}{L}, \text{ and}$$

$$C = 2b + \underset{\sim}{\mu}'Q\underset{\sim}{\mu} + \sum_{1}^{n} y_t^2 + y_{n+1}^2.$$

Then (4.4.1) can be written as

$$p(y_1,y_2,\ldots,y_{n+1},\underset{\sim}{\phi},\tau) \propto \tau^{\frac{n+2a+q+1}{2}-1} \exp -\frac{\tau}{2}[(\underset{\sim}{\phi}-G^{-1}\underset{\sim}{Z})'G(\underset{\sim}{\phi}-G^{-1}\underset{\sim}{Z}) + C-\underset{\sim}{Z}'G^{-1}\underset{\sim}{Z}]. \tag{4.4.2}$$

The joint distribution of $y_1,y_2,\ldots,y_{n+1}$ is

$$p(y_1,y_2,\ldots,y_{n+1}) = \int_\tau \int_{\underset{\sim}{\phi}} p(y_1,y_2,\ldots,y_{n+1},\underset{\sim}{\phi},\tau)\,d\underset{\sim}{\phi}\,d\tau$$

$$\propto \int_0^\infty \tau^{\frac{n+2a+q+1}{2}-1} e^{-\frac{\tau}{2}[C-\underset{\sim}{Z}'G^{-1}\underset{\sim}{Z}]} \cdot \frac{1}{\tau^{q/2}\,|G|^{1/2}}\,d\tau$$

$$\propto \frac{1}{|G|^{1/2}\,[C-\underset{\sim}{Z}'G^{-1}\underset{\sim}{Z}]^{\frac{n+2a+1}{2}}}. \tag{4.4.3}$$

Let

$$D = 1 - \underset{\sim}{L}'G^{-1}\underset{\sim}{L},$$

$$E = -\underset{\sim}{L}'G^{-1}\underset{\sim}{B}, \text{ and}$$

$$F = 2b + \underset{\sim}{\mu}'Q\underset{\sim}{\mu} + \sum_1^n y_t^2 - \underset{\sim}{B}G^{-1}\underset{\sim}{B} .$$

Thus (4.4.3) can be written as

$$p(y_1,y_2,\ldots,y_{n+1}) \propto \frac{1}{|G|^{1/2}\,[(y_{n+1} - E/D)^2 D + F - E^2/D]^{\frac{n+2a+1}{2}}} .$$

So, the predictive density of y_{n+1} is

$$p(y_{n+1}|\underset{\sim}{Y}) \propto \frac{1}{[(y_{n+1} - E/D)^2 D + F - E^2/D]^{\frac{n+2a+1}{2}}} , \quad y_{n+1} \in R,\ \underset{\sim}{y} \in R^n \quad (4.4.4)$$

i.e., the predictive density of the next observation is t with $n+2a$ degrees of freedom, E/D location parameter, and $\frac{D(n+2a)}{F - E^2/D}$ precision.

4.5 Conclusions and Comments

The main objective of this chapter was to develop a convenient and adequate approximation to the posterior distributions of the parameters of the moving average process. The methodology introduced here is based on equating the initial values of the errors to zeroes. It has been shown that Bayesian inferences about the parameters can be approximately done using a t distribution. Based on the t approximation, inferences about the parameters $\underset{\sim}{\phi}$ can be done without the need of using any numerical techniques which saves money and computing time. Highest posterior density (H.P.D.) regions for the parameters can be constructed, and a Bayesian way to estimate the order of the processes can be found. Furthermore, the use of t approximation enables us to obtain marginal distributions for the next observation and the error precision which was impossible to obtain even for the first order process.

CHAPTER V

SOME NUMERICAL STUDIES

To demonstrate the idea of using the exact conditional likelihood function and the t approximation discussed in Chapter IV, two numerical studies will be discussed. The first one deals with generating five data sets from a MA(1) model with parameter values -1, -.5, 0, .5, and 1 for the moving average coefficient. The Bayesian approach is developed to analyze the data sets using the exact conditional likelihood function and the t approximation. The second numerical study deals with generating four data sets from a MA(2) model with parameters (-.5,0), (-.5,.5), (.3,.1), and (.5,.5). The Bayesian approach is developed to analyze the data sets using the exact conditional likelihood function and the t approximation.

The two numerical examples are also analyzed using Box and Jenkins' technique.

5.1 The Numerical Study of a MA(1) Process

As we have mentioned, the Bayesian approach has been used in two different ways to analyze this numerical study. The first way uses the idea of setting the initial values of the errors to zeroes in evaluating the likelihood function. The other way uses the t approximation introduced in (4.2) to make the inferences about the parameters. For the same data sets, Box and Jenkins' procedure is used to make inferences about the model parameters.

5.1.1 A Numerical Study Using the Exact Likelihood Function

Let q = 1 in (4.1.7), then the marginal posterior distribution $\xi(\phi_1|\underset{\sim}{Y})$ can be rewritten as

$$\xi(\phi_1|\underset{\sim}{Y}) \propto \frac{1}{\left[2b + Q(\phi_1-\mu)^2 + \sum_{t=1}^{n}(y_t + \phi_1\varepsilon_{t-1}(\phi_1))^2\right]^{\frac{n+2a+1}{2}}}, \tag{5.1.1.1}$$

$$\phi_1 \in R, \quad b > 0, \quad a > 0, \quad \mu \in R, \quad \underset{\sim}{y} \in R^n.$$

The five data which have been generated from a MA(1) process have different values for ϕ_1 and a common value of τ. The chosen values of ϕ_1 are -1, -.5, 0, .5, and 1; the chosen value of τ is 2; while the chosen value of the hyperparameter vector $(a,b,\phi,Q)^T$ is $(-\frac{1}{2},0,\phi_1,0)$. Sample sizes 30, 50, 70, and 90 are chosen to see the effect of adding more observations. The posterior expectation, variance, and coefficient of variation are calculated using the Gaussian-quadrature method of numerical integration for each data set. The results are shown in Tables I, II, III, IV, and V. The expectation is denoted by $E(\phi_1|Y)$, the variance is denoted by $V(\phi_1|\underset{\sim}{Y})$, and the coefficient of variation is denoted by $CV(\phi_1|\underset{\sim}{Y})$.

The numerical results are consistent with the theoretical analysis. The posterior expectation, as an estimate for ϕ_1, becomes closer to the true values as n increases; the precision in estimating ϕ_1 by the posterior expectation is improved by adding more observations.

5.1.2 The Numerical Study Using the t Approximation

Here, τ and ϕ_1 are assumed to be independent apriori and have vague prior distributions. The joint prior distribution of τ and ϕ_1 can be obtained from (4.1.5) by setting $q = 1$, $Q = 0$, $b = 0$, and $a = -\frac{1}{2}$. Thus, the marginal posterior distribution of ϕ_1 is a univariate t with $n-1$ degrees of freedom, B/A a location parameter, and $\frac{(n-1)A}{C - B^2/A}$ precision where A, B, and C are defined as the same as they were used in section 4.2 with the above hyperparameter values.

Thus,

$$\phi_1 | \underset{\sim}{Y} \sim T_1 (n-1,\ B/A,\ \frac{(n-1)\ A}{C - B^2/A}) \tag{5.1.2.1}$$

where

$$A = \sum_{t=1}^{n} \varepsilon_{t-1}^2 (\underset{\sim}{\phi}_0),$$

$$B = -\sum_{1}^{n} y_t \varepsilon_{t-1} (\underset{\sim}{\phi}_0), \text{ and}$$

$$C = \sum_{1}^{n} y_t^2.$$

Hence, the expectation and variance of $\phi_1 | \underset{\sim}{Y}$ are B/A and $\frac{(n-1)(C - B^2/A)}{(n-1)(n-3)}$, respectively. Similarly,

$$Y_{n+1} | \underset{\sim}{Y} \sim T_1 (n - 1,\ E/D,\ \frac{(n-1) D}{F - E^2/D}) \tag{5.1.2.2}$$

where

$$D = 1 - \varepsilon_n^2 (\phi_{10})/A,$$

$$E = -\varepsilon_n (\phi_{10}) B/A, \quad \text{and}$$

$$F = \sum_{1}^{n} y_t^2 - B^2/A.$$

Thus, the expectation and variance of $Y_{n+1}|Y$ are E/D and $\frac{(n-1)(F - E^2/D)}{(n-1)(n-3)D}$, respectively. The same generated data sets have been analyzed using the above t approximation for the same sample sizes 30, 50, 70, and 90. The posterior expectation, variance, and coefficient of variation of ϕ_1, and also the expectation and variance of Y_{n+1} are calculated using the above results. The results are shown in Tables VI, VII, VIII, IX, and X.

The expectation, variance, and coefficient of variation of ϕ_1 are denoted by $E(\phi_1|\underset{\sim}{Y})$, $V(\phi_1|\underset{\sim}{Y})$, and $CV(\phi_1|\underset{\sim}{Y})$, respectively. Also the expectation and variance of Y_{n+1} are denoted by $E(Y_{n+1}|\underset{\sim}{Y})$ and $V(Y_{n+1}|\underset{\sim}{Y})$, respectively.

Inspection of the numerical results given in the Tables VI, VII, VIII, IX, and X supports the adequacy of the t approximation in analyzing the MA(1) model. The posterior expectation $E(\phi_1|\underset{\sim}{Y})$, as an estimate for ϕ_1, becomes closer to the true parameter as the sample size increases. The precision in estimating ϕ_1 is improved by adding more observations. We also notice that the posterior expectation and variance become stable with adding more observations when the sample size reaches 70.

5.1.3 The Numerical Study Using Box and Jenkins Procedure

The methodology introduced by Box and Jenkins (1970) is essentially classical and restricted by the invertibility conditions on the parameters space. This methodology is used here to analyze the same previous data sets for the same sample sizes 30, 50, 70, and 90. The results are shown in Tables XI, XII, XIII, XIV, and XV. The estimate of ϕ_1 is denoted by

$\hat{\phi}_1$, the variance of $\hat{\phi}_1$ is denoted by $V(\hat{\phi}_1)$, the coefficient of variation of $\hat{\phi}_1$ is denoted by $CV(\hat{\phi}_1)$, the estimate of Y_{n+1} is denoted by $\hat{Y}_{n+1}$, and the variance of $\hat{Y}_{n+1}$ is denoted by $V(\hat{Y}_{n+1})$.

It is important to notice that the generated process is not invertible at $\phi_1 = -1$ and $\phi_1 = 1$. Actually these are on the boundaries of the invertibility domain of the parameter ϕ_1. However, Box and Jenkins' procedure converges for $n = 30$ and $\phi_1 = 1$ after 49 iterations.

It is also important to mention that all computations here were done on the ETS package of SAS. The MA(1) model used by ETS is defined by $Y_t = \mu + \varepsilon_t - \phi_1 \varepsilon_{t-1}$ where $t = 1,2,\ldots,n$ and $\mu \in R$. However, most of the estimates calculated here match the corresponding estimates obtained by using the t approximation especially when the process is invertible and the sample size is large.

5.2 The Numerical Study of a MA(2) Process

The four data sets which have been generated from a MA(2) process have different values for (ϕ_1,ϕ_2) and a fixed value of τ. The chosen combinations of (ϕ_1,ϕ_2) are $(-.5,0)$, $(-.5,.5)$, $(.3,.1)$, and $(.5,-.5)$; the chosen value of τ is 2. These data sets are analyzed using the same three procedures which have been used to analyze the generated data sets of the MA(1) study.

5.2.1 The Exact Conditional Likelihood Function Procedure

In this analysis, the prior distribution of τ is assumed to be gamma, the prior distribution of (ϕ_1,ϕ_2) is assumed to be vague, and τ and (ϕ_1,ϕ_2) are assumed to be independent. The posterior distribution of

(ϕ_1,ϕ_2) can be obtained from (4.1.7) as follows

$$\xi(\phi_1,\phi_2|\underset{\sim}{Y}) \propto \frac{1}{[2b + \sum_{t=1}^{n} (Y_t + \phi_1\varepsilon_{t-1}(\underset{\sim}{\phi}) + \phi_2\varepsilon_{t-2}(\underset{\sim}{\phi}))^2]^{\frac{n+2a}{2}}}, \qquad (5.2.1.1)$$

$$(\phi_1,\phi_2) \in R^2, \quad a > 0, \quad b > 0, \quad Y \in R^n.$$

The chosen value of the hyperparameter vector $(a,b)^T$ is $(-1,0)^T$. The sample sizes 30, 50, 70, and 90 are kept the same as were used with MA(1) process. The posterior expectations, variances, and the coefficients of variations of ϕ_1 and ϕ_2 are calculated numerically using the Gaussian-quadrature method in performing the integration processes. Furthermore, the posterior conditional expectation and variance of the predictive density of the one-step-ahead forecast given the estimated values of ϕ_1 and ϕ_2, and the correlation coefficient of ϕ_1 and ϕ_2 are calculated. The numerical results are shown in Tables XVI, XVII, XVIII, and XIX. The posterior expectation, variance, and coefficient of variation of ϕ_1 are denoted by $E(\phi_1|\underset{\sim}{Y})$, $V(\phi_1|\underset{\sim}{Y})$, and $CV(\phi_1|\underset{\sim}{Y})$, respectively. Similarly the posterior expectation, variance, and coefficient of variation of ϕ_2 are denoted by $E(\phi_2|\underset{\sim}{Y})$, $V(\phi_2|\underset{\sim}{Y})$, and $CV(\phi_2|\underset{\sim}{Y})$, respectively. The correlation coefficient of ϕ_1 and ϕ_2 is denoted by $\rho(\phi_1,\phi_2|\underset{\sim}{Y})$. Also, the conditional expectation and variance of the next observation given the estimates of ϕ_1 and ϕ_2 are denoted by $E(Y_{n+1}|\underset{\sim}{\phi})$ and $V(Y_{n+1}|\underset{\sim}{\phi})$, respectively.

From the numerical results, it is easy to see that the posterior expectations vector $(E(\phi_1|\underset{\sim}{Y}),E(\phi_2|\underset{\sim}{Y}))^T$, as an estimate for the parameters vector $\underset{\sim}{\phi} = (\phi_1,\phi_2)^T$, becomes closer to the true vector as the sample size increases. The precision in estimating ϕ_1 and ϕ_2 is improved by adding more observations.

5.2.2 The Analysis of the Data Using the t-Approximation

In this analysis, τ and $\underset{\sim}{\phi}$ are assumed to be independent apriori and prior distributions. The posterior distribution of $(\phi_1, \phi_2)^T$ can be obtained from (4.2.1) by putting a = -1, b = 0 and $Q_{2\times 2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Thus,

$$\underset{\sim}{\phi} \sim T_2(n-2, A^{-1}\underset{\sim}{B}, \frac{A(n-2)}{C - \underset{\sim}{B}'A^{-1}B})$$

where

$$A_{2\times 2} = (a_{ij}) = (\sum_{t=1}^{n} \varepsilon_{t-i}(\underset{\sim}{\phi}_0)\varepsilon_{t-j}(\underset{\sim}{\phi}_0)), \qquad i,j = 1,2,$$

$$\underset{\sim}{B}_{2\times 1} = (b_j) = (\sum_{t=1}^{n} Y_t \varepsilon_{t-j}(\underset{\sim}{\phi}_0)), \qquad j = 1,2, \text{ and}$$

$$C = \sum_1^n Y_t^2.$$

Thus, the expectation and variance of $\underset{\sim}{\phi}|\underset{\sim}{Y}$ are

$$A^{-1}\underset{\sim}{B} \quad \text{and} \quad \frac{(n-2)(C - \underset{\sim}{B}'A^{-1}\underset{\sim}{B})A^{-1}}{(n-2)(n-4)}, \text{ respectively.} \tag{5.2.2.1}$$

Similarly, $Y_{n+1}|\underset{\sim}{Y} \sim T_1(n-2, D^{-1}\underset{\sim}{E}, \frac{(n-2)D}{F - \underset{\sim}{E}'D^{-1}\underset{\sim}{E}})$ where D, $\underset{\sim}{E}$, and F are defined as in section 4.4 for q = 2.

Thus the expectation and variance of $Y_{n+1}|\underset{\sim}{Y}$ are

$$D^{-1}\underset{\sim}{E} \quad \text{and} \quad \frac{(n-2)(F - \underset{\sim}{E}'D^{-1}\underset{\sim}{E})D^{-1}}{(n-2)(n-4)}, \text{ respectively.} \tag{5.2.2.2}$$

The same data sets generated from a MA(2) model have been analyzed using the t approximation for the same sample sizes 30, 50, 70, and 90. The posterior expectations, variances, coefficients of variation, and the correlation coefficient of ϕ_1 and ϕ_2 are calculated using (5.2.2.1). The marginal expectation and variance of Y_{n+1} are also calculated using (5.2.2.2). The results are shown in Tables XX, XXI, XXII, and XIII. The symbols $E(\phi_1|\underset{\sim}{Y})$, $V(\phi_1|\underset{\sim}{Y})$, $CV(\phi_1|\underset{\sim}{Y})$, $E(\phi_2|\underset{\sim}{Y})$, $V(\phi_2|\underset{\sim}{Y})$, $CV(\phi_2|\underset{\sim}{Y})$ and $\rho(\phi_1,\phi_2|\underset{\sim}{Y})$ have the same meaning as they were used in section 5.2.1. The marginal expectation and variance of Y_{n+1} are denoted by $E(Y_{n+1}|\underset{\sim}{Y})$ and $V(Y_{n+1}|\underset{\sim}{Y})$, respectively.

Inspectation of the numerical results supports the adequacy of the t approximation in analyzing the MA(2) model. The results are consistent with the theoretical objectives; the posterior expectations vector $(E(\phi_1|\underset{\sim}{Y}), E(\phi_2|\underset{\sim}{Y}))^T$ becomes closer to the true parameters vector $(\phi_1\ \phi_2)^T$ as n increases; the precision in estimating ϕ_1 and ϕ_2 is improved by adding more observations.

5.2.3 The Analysis of the Data Using the Box and Jenkins' Procedure

Box and Jenkins' technique is used again to analyze the same previous data sets generated from a MA(2) model for the same sample sizes 30, 50, 70, and 90. The results are shown in Tables XXIV, XXV, XXVI, and XXVII. The estimate of ϕ_1 is denoted by $\hat{\phi}_1$, the variance of $\hat{\phi}_1$ is denoted by $V(\hat{\phi}_1)$, and the coefficient of variation of $\hat{\phi}_1$ is denoted by $CV(\hat{\phi}_1)$. Similarly, the estimate of ϕ_2, the variance, and the coefficient of variation are denoted by $\hat{\phi}_2$, $V(\hat{\phi}_2)$, and $CV(\hat{\phi}_2)$, respectively. Also, the estimate of Y_{n+1} is denoted by $\hat{Y}_{n+1}$, the variance of $\hat{Y}_{n+1}$ is denoted by $V(\hat{Y}_{n+1})$, and the correlation coefficient of $\hat{\phi}_1$ and $\hat{\phi}_2$ is denoted by $\rho(\hat{\phi}_1,\hat{\phi}_2)$.

Again, all computations here were done on ETS package of SAS. The MA(2) model used by ETS is defined by $Y_t = \mu + \varepsilon_t - \phi_1\varepsilon_{t-1} - \phi_2\varepsilon_{t-2}$ where $t = 1,2,\ldots,n$ and $\mu \in R$. However, we can see that most of the estimates calculated here match the corresponding estimates obtained by using the t approximation when n is larger and the process is invertible.

5.3 Conclusions

The main objective of the numerical studies is to show how one can use the two Bayesian procedures introduced in Chapter IV in analyzing the moving average process and to see the adequacy of using the t approximation in making inferences about the parameters. To do that, the data sets gneerated from a MA(1) and a MA(2) process have been analyzed using the three different procedures to see how much the estimates are far from the true values. Inspection of the numerical results supports the adequacy of using the t approximation in analyzing moving average processes.

Based on t approximation, one can develop a complete Bayesian theory of MA(q) analysis. The t approximation procedure gives a way by which H.P.D. regions can be constructed and hence gives a Bayesian technique to estimate the order of the process. Moreover, the t approximation is easy to use and economical. Using the t approximation, one may also develop a complete Bayesian theory to analyze ARMA(p,q) processes. Although the numerical results obtained by the t approximation about the parameters match the results obtained by Box and Jenkins' procedure, it seems that the convergence of the t approximation is slower than the convergence of the exact procedure as expected.

TABLE I

MODERATE SAMPLE BEHAVIOR OF $P(\phi_1 | \underset{\sim}{Y})$
USING THE EXACT PROCEDURE
FOR $\phi_1 = -1$

N	$E(\phi_1 \| \underset{\sim}{Y})$	$V(\phi_1 \| \underset{\sim}{Y})$	$CV(\phi_1 \| \underset{\sim}{Y})$
30	-0.7847	0.0278	-0.2125
50	-0.8645	0.0083	-0.1054
70	-0.9147	0.0048	-0.0757
90	-0.9363	0.0032	-0.0604

TABLE II

MODERATE SAMPLE BEHAVIOR OF $P(\phi_1 | \underset{\sim}{Y})$ USING THE EXACT PROCEDURE FOR $\phi_1 = -.5$

N	$E(\phi_1 \| \underset{\sim}{Y})$	$V(\phi_1 \| \underset{\sim}{Y})$	$CV(\phi_1 \| \underset{\sim}{Y})$
30	-0.2975	0.0190	-0.4634
50	-0.3610	0.0125	-0.3099
70	-0.4343	0.0081	-0.2070
90	-0.4537	0.0063	-0.1748

TABLE III

MODERATE SAMPLE BEHAVIOR OF $P(\phi_1|\underset{\sim}{Y})$
USING THE EXACT PROCEDURE
FOR $\phi_1 = 0$

N	$E(\phi_1\|\underset{\sim}{Y})$	$V(\phi_1\|\underset{\sim}{Y})$	$CV(\phi_1\|\underset{\sim}{Y})$
30	0.0855	0.0173	1.5391
50	0.0665	0.0134	1.7402
70	-0.0209	0.0092	-4.5843
90	-0.0296	0.0075	-2.9160

TABLE IV

MODERATE SAMPLE BEHAVIOR OF $P(\phi_1 | \underset{\sim}{Y})$
USING THE EXACT PROCEDURE
FOR $\phi_1 = .5$

N	$E(\phi_1 \mid \underset{\sim}{Y})$	$V(\phi_1 \mid \underset{\sim}{Y})$	$CV(\phi_1 \mid \underset{\sim}{Y})$
30	0.4491	0.0144	0.2669
50	0.4758	0.0108	0.2189
70	0.3938	0.0085	0.2339
90	0.4000	0.0064	0.2006

TABLE V

MODERATE SAMPLE BEHAVIOR OF $P(\phi_1 | \underset{\sim}{Y})$
USING THE EXACT PROCEDURE
FOR $\phi_1 = 1$

N	$E(\phi_1 \mid \underset{\sim}{Y})$	$V(\phi_1 \mid \underset{\sim}{Y})$	$CV(\phi_1 \mid \underset{\sim}{Y})$
30	1.0790	0.0170	0.1207
50	0.9219	0.0053	0.0792
70	0.8774	0.0052	0.0822
90	0.9032	0.0047	0.0757

TABLE VI

MODERATE SAMPLE BEHAVIOR OF $P(\phi_1|\underset{\sim}{Y})$ USING THE t APPROXIMATION FOR $\phi_1 = -1$

N	$E(\phi_1\|\underset{\sim}{Y})$	$V(\phi_1\|\underset{\sim}{Y})$	$CV(\phi_1\|\underset{\sim}{Y})$	$E(Y_{n+1}\|\underset{\sim}{Y})$	$V(Y_{n+1}\|\underset{\sim}{Y})$
30	-0.8623	0.0396	-0.2308	-0.9335	0.6335
50	-0.9250	0.0213	-0.1577	0.1216	0.5868
70	-1.0328	0.0151	-0.1191	-0.7326	0.5526
90	-1.0487	0.0114	-0.1020	-0.4292	0.5387

TABLE VII

MODERATE SAMPLE BEHAVIOR OF $P(\phi_1|\underset{\sim}{Y})$ USING THE t APPROXIMATION FOR $\phi_1 = -.5$

N	$E(\phi_1\|\underset{\sim}{Y})$	$V(\phi_1\|\underset{\sim}{Y})$	$CV(\phi_1\|\underset{\sim}{Y})$	$E(Y_{n+1}\|\underset{\sim}{Y})$	$V(Y_{n+1}\|\underset{\sim}{Y})$
30	-0.4069	0.0385	-0.4822	-0.3344	0.5942
50	-0.4388	0.0212	-0.3314	-0.3858	0.5818
70	-0.5503	0.0149	-0.2222	-0.3928	0.5433
90	-0.5503	0.0115	-0.1946	-0.2850	0.5415

TABLE VIII

MODERATE SAMPLE BEHAVIOR OF $P(\phi_1|\underset{\sim}{Y})$ USING THE t APPROXIMATION FOR $\phi_1 = 0$

N	$E(\phi_1\|\underset{\sim}{Y})$	$V(\phi_1\|\underset{\sim}{Y})$	$CV(\phi_1\|\underset{\sim}{Y})$	$E(Y_{n+1}\|\underset{\sim}{Y})$	$V(Y_{n+1}\|\underset{\sim}{Y})$
30	0.1374	0.0388	1.4334	0.1208	0.6316
50	0.0792	0.0213	0.1841	0.0088	0.5979
70	-0.0220	0.0151	-5.5829	-0.0139	0.5551
90	-0.0333	0.0116	-3.2329	-0.0210	0.5486

TABLE IX

MODERATE SAMPLE BEHAVIOR OF $P(\phi_1|\underset{\sim}{Y})$ USING THE
t APPROXIMATION FOR $\phi_1 = .5$

N	$E(\phi_1\|\underset{\sim}{Y})$	$V(\phi_1\|\underset{\sim}{Y})$	$CV(\phi_1\|\underset{\sim}{Y})$	$E(Y_{n+1}\|\underset{\sim}{Y})$	$V(Y_{n+1}\|\underset{\sim}{Y})$
30	0.6938	0.0372	0.2780	0.6895	0.6155
50	0.6003	0.0209	0.2409	0.0404	0.5921
70	0.4894	0.0150	0.2503	0.2722	0.5402
90	0.4793	0.0116	0.2247	0.3433	0.5417

TABLE X

MODERATE SAMPLE BEHAVIOR OF $P(\phi_1 | \underset{\sim}{Y})$ USING THE t APPROXIMATION FOR $\phi_1 = 1$

N	$E(\phi_1 \mid \underset{\sim}{Y})$	$V(\phi_1 \mid \underset{\sim}{Y})$	$CV(\phi_1 \mid \underset{\sim}{Y})$	$E(Y_{n+1} \mid \underset{\sim}{Y})$	$V(Y_{n+1} \mid \underset{\sim}{Y})$
30	1.5254	0.0427	0.1354	2.6764	0.5915
50	1.1127	0.0212	0.1309	-0.2654	0.5888
70	1.0459	0.0150	0.1169	0.2756	0.5031
90	0.9664	0.0116	0.1113	0.6047	0.5446

TABLE XI

MODERATE SAMPLE BEHAVIOR OF BOX AND JENKINS PROCEDURE FOR $\phi_1 = -1$

N	$\hat{\phi}_1$	$V(\hat{\phi}_1)$	$CV(\hat{\phi}_1)$	$\hat{Y}_{n+1}$	$V(\hat{Y}_{n+1})$
30	-0.6295	0.0251	-0.2515	-0.1804	0.5226
50	-0.9000	0.0041	-0.0711	0.1718	0.5721
70	-0.9301	0.0020	-0.0481	-0.7204	0.5351
90	-0.9477	0.0013	-0.0378	-0.4226	0.5310

TABLE XII

MODERATE SAMPLE BEHAVIOR OF BOX AND JENKINS PROCEDURE FOR $\phi_1 = -.5$

N	$\hat{\phi}_1$	$V(\hat{\phi}_1)$	$CV(\hat{\phi}_1)$	$\hat{Y}_{n+1}$	$V(\hat{Y}_{n+1})$
30	-0.2031	0.0379	-0.9585	0.1291	0.4757
50	-0.3714	0.0180	-0.3609	0.0330	0.5684
70	-0.4484	0.0118	-0.2424	-0.3706	0.5311
90	-0.4651	0.0091	-0.2054	-0.2945	0.5341

TABLE XIII

MODERATE SAMPLE BEHAVIOR OF BOX AND JENKINS PROCEDURE FOR $\phi_1 = 0$

N	$\hat{\phi}_1$	$V(\hat{\phi}_1)$	$CV(\hat{\phi}_1)$	$\hat{Y}_{n+1}$	$V(\hat{Y}_{n+1})$
30	0.2072	0.0366	0.9228	0.4544	0.5042
50	0.0756	0.0207	1.9031	0.0802	0.5797
70	-0.0194	0.0148	-6.2826	-0.0580	0.5351
90	-0.0280	0.0114	-3.8146	-0.0610	0.5358

TABLE XIV

MODERATE SAMPLE BEHAVIOR OF BOX AND JENKINS PROCEDURE FOR $\phi_1 = .5$

N	$\hat{\phi}_1$	$V(\hat{\phi}_1)$	$CV(\hat{\phi}_1)$	$\hat{Y}_{n+1}$	$V(\hat{Y}_{n+1})$
30	0.6032	0.0230	0.2514	0.7775	0.5232
50	0.5088	0.0156	0.2454	0.1158	0.5842
70	0.4069	0.0126	0.2758	0.1714	0.5271
90	0.4128	0.0094	0.2349	0.2718	0.5291

TABLE XV

MODERATE SAMPLE BEHAVIOR OF BOX AND JENKINS PROCEDURE FOR $\phi_1 = 1$

N	$\hat{\phi}_1$	$V(\hat{\phi}_1)$	$CV(\hat{\phi}_1)$	$\hat{Y}_{n+1}$	$V(\hat{Y}_{n+1})$
30	1.3269	0.0001	0.0066	1.0624	0.3593
50	1.1804	0.0000	0.0030	-0.2005	0.4416
70	0.9621	0.0020	0.0465	0.2796	0.5096
90	0.9255	0.0021	0.0493	0.5472	0.5351

TABLE XVI

MODERATE SAMPLE BEHAVIOR OF $P(\phi_1,\phi_2|\underset{\sim}{Y})$ USING THE EXACT PROCEDURE FOR $\phi_1 = -.5$ AND $\phi_2 = .5$

N	$E(\phi_1\|\underset{\sim}{Y})$	$V(\phi_1\|\underset{\sim}{Y})$	$CV(\phi_1\|\underset{\sim}{Y})$	$E(\phi_2\|\underset{\sim}{Y})$	$V(\phi_2\|\underset{\sim}{Y})$	$CV(\phi_2\|\underset{\sim}{Y})$	$\rho(\phi_1,\phi_2\|\underset{\sim}{Y})$	$E(Y_{n+1}\|\underset{\sim}{\phi})$	$V(Y_{n+1}\|\underset{\sim}{\phi})$
30	-0.5458	0.0416	-0.3735	0.2962	0.0414	0.6868	-0.1020	0.2070	0.5780
50	-0.4847	0.0330	-0.3750	0.3125	0.0165	0.4112	0.4470	0.5570	0.5750
70	-0.5374	0.0078	-0.1638	0.3694	0.0106	0.2780	0.1240	-0.1820	0.5160
90	-0.5217	0.0031	-0.1059	0.4281	0.0117	0.2528	0.0010	-0.1940	0.5270

TABLE XVII

MODERATE SAMPLE BEHAVIOR OF $P(\phi_1,\phi_2|\underset{\sim}{Y})$ USING THE EXACT PROCEDURE FOR $\phi_1 = -.5$ AND $\phi_2 = 0.0$

N	$E(\phi_1\|\underset{\sim}{Y})$	$V(\phi_1\|\underset{\sim}{Y})$	$CV(\phi_1\|\underset{\sim}{Y})$	$E(\phi_2\|\underset{\sim}{Y})$	$V(\phi_2\|\underset{\sim}{Y})$	$CV(\phi_2\|\underset{\sim}{Y})$	$\rho(\phi_1,\phi_2\|\underset{\sim}{Y})$	$E(Y_{n+1}\|\underset{\sim}{\phi})$	$V(Y_{n+1}\|\underset{\sim}{\phi})$
30	-0.3971	0.0291	-0.4294	-0.3214	0.0346	-0.5790	0.0500	0.8150	0.5400
50	-0.4147	0.0192	-0.3341	-0.1472	0.0162	-0.8638	0.1830	0.3110	0.5600
70	-0.4918	0.0135	-0.2360	-0.1861	0.0110	-0.5643	0.0850	0.1290	0.5060
90	-0.5223	0.0100	-0.1912	-0.1380	0.0069	-0.6021	0.3330	0.1070	0.5160

TABLE XVIII

MODERATE SAMPLE BEHAVIOR OF $P(\phi_1,\phi_2|\underset{\sim}{Y})$ USING THE EXACT PROCEDURE FOR $\phi_1 = .3$ AND $\phi_2 = .1$

N	$E(\phi_1\|\underset{\sim}{Y})$	$V(\phi_1\|\underset{\sim}{Y})$	$CV(\phi_1\|\underset{\sim}{Y})$	$E(\phi_2\|\underset{\sim}{Y})$	$V(\phi_2\|\underset{\sim}{Y})$	$CV(\phi_2\|\underset{\sim}{Y})$	$\rho(\phi_1,\phi_2\|\underset{\sim}{Y})$	$E(Y_{n+1}\|\underset{\sim}{\phi})$	$V(Y_{n+1}\|\underset{\sim}{\phi})$
30	0.2670	0.0349	0.6996	-0.3312	0.0421	-0.6198	0.1370	0.0450	0.5520
50	0.3347	0.0194	0.4162	-0.1036	0.0169	-1.2543	-0.0980	-0.2300	0.5730
70	0.2581	0.0134	0.4490	-0.1451	0.0124	-0.7675	-0.1230	0.0890	0.5020
90	0.2660	0.0100	0.3752	-0.0611	0.0067	-1.3418	-0.2450	-0.0010	0.5150

TABLE XIX

MODERATE SAMPLE BEHAVIOR OF $P(\phi_1,\phi_2|\underset{\sim}{Y})$ USING THE EXACT PROCEDURE FOR $\phi_1 = .5$ AND $\phi_2 = -.5$

N	$E(\phi_1\|\underset{\sim}{Y})$	$V(\phi_1\|\underset{\sim}{Y})$	$CV(\phi_1\|\underset{\sim}{Y})$	$E(\phi_2\|\underset{\sim}{Y})$	$V(\phi_2\|\underset{\sim}{Y})$	$CV(\phi_2\|\underset{\sim}{Y})$	$\rho(\phi_1,\phi_2\|\underset{\sim}{Y})$	$E(Y_{n+1}\|\underset{\sim}{\phi})$	$V(Y_{n+1}\|\underset{\sim}{\phi})$
30	0.4101	0.0452	0.5186	-0.7958	0.0279	-0.2099	0.2330	0.5130	0.5800
50	0.4760	0.0168	0.2721	-0.5513	0.0096	-0.1779	-0.0490	-0.3100	0.5910
70	0.4064	0.0146	0.2970	-0.5695	0.0088	-0.1648	0.1610	0.3210	0.5220
90	0.4756	0.0093	0.2027	-0.5377	0.0025	-0.0939	-0.2390	0.3650	0.5320

TABLE XX

MODERATE SAMPLE BEHAVIOR OF $P(\phi_1,\phi_2|\underset{\sim}{Y})$ USING THE t APPROXIMATION FOR $\phi_1 = -.5$ AND $\phi_2 = .5$

N	$E(\phi_1\|\underset{\sim}{Y})$	$V(\phi_1\|\underset{\sim}{Y})$	$CV(\phi_1\|\underset{\sim}{Y})$	$E(\phi_2\|\underset{\sim}{Y})$	$V(\phi_2\|\underset{\sim}{Y})$	$CV(\phi_2\|\underset{\sim}{Y})$	$\rho(\phi_1,\phi_2\|\underset{\sim}{Y})$	$E(Y_{n+1}\|\underset{\sim}{Y})$	$V(Y_{n+1}\|\underset{\sim}{Y})$
30	-0.4970	0.0360	-0.3818	0.2230	0.0390	0.8856	-0.0170	-0.2300	0.6600
50	-0.4150	0.0220	-0.3574	0.3150	0.0230	0.4815	-0.0020	-0.4900	0.6300
70	-0.5200	0.0149	-0.2347	0.2340	0.0150	0.5234	-0.0200	0.1480	0.5278
90	-0.5360	0.0116	-0.2009	0.3060	0.0117	0.3535	-0.0002	0.0370	0.5310

TABLE XXI

MODERATE SAMPLE BEHAVIOR OF $P(\phi_1,\phi_2|\underset{\sim}{Y})$ USING THE t APPROXIMATION FOR $\phi_1 = -.5$ AND $\phi_2 = 0$

N	$E(\phi_1\|\underset{\sim}{Y})$	$V(\phi_1\|\underset{\sim}{Y})$	$CV(\phi_1\|\underset{\sim}{Y})$	$E(\phi_2\|\underset{\sim}{Y})$	$V(\phi_2\|\underset{\sim}{Y})$	$CV(\phi_2\|\underset{\sim}{Y})$	$\rho(\phi_1,\phi_2\|\underset{\sim}{Y})$	$E(Y_{n+1}\|\underset{\sim}{Y})$	$V(Y_{n+1}\|\underset{\sim}{Y})$
30	-0.4854	0.0396	-0.4100	-0.3677	0.0423	-0.5595	0.0234	-0.9585	0.6453
50	-0.4501	0.0212	-0.3237	-0.1553	0.0212	-0.9382	-0.0300	-0.3491	0.5762
70	-0.5453	0.0146	-0.2214	-0.2568	0.0147	-0.4725	-0.0273	-0.1764	0.5068
90	-0.5511	0.0113	-0.1930	-0.1996	0.0114	-0.5359	-0.0121	-0.1366	0.5202

TABLE XXII

MODERATE SAMPLE BEHAVIOR OF $P(\phi_1,\phi_2|\underset{\sim}{Y})$ USING THE t APPROXIMATION FOR $\phi_1 = .3$ AND $\phi_2 = -.1$

N	$E(\phi_1\|Y)$	$V(\phi_1\|Y)$	$CV(\phi_1\|Y)$	$E(\phi_2\|Y)$	$V(\phi_2\|Y)$	$CV(\phi_2\|Y)$	$\rho(\phi_1,\phi_2\|Y)$	$E(Y_{n+1}\|Y)$	$V(Y_{n+1}\|Y)$
30	0.3256	0.0469	0.6651	-0.3100	0.0452	-0.6858	0.1914	0.0532	0.7083
50	0.3618	0.0212	0.4025	-0.1098	0.0212	-1.3260	0.0431	0.2507	0.5885
70	0.2749	0.0147	0.4406	-0.1477	0.0148	-0.8223	0.0145	-0.0885	0.5046
90	0.2753	0.0114	0.3878	-0.0742	0.0115	-1.4453	-0.0005	0.0039	0.5239

TABLE XXIII

MODERATE SAMPLE BEHAVIOR OF $P(\phi_1,\phi_2|\underset{\sim}{Y})$ USING THE t APPROXIMATION FOR $\phi_1 = .5$ AND $\phi_2 = -.5$

N	$E(\phi_1\|\underset{\sim}{Y})$	$V(\phi_1\|\underset{\sim}{Y})$	$CV(\phi_1\|\underset{\sim}{Y})$	$E(\phi_2\|\underset{\sim}{Y})$	$V(\phi_2\|\underset{\sim}{Y})$	$CV(\phi_2\|\underset{\sim}{Y})$	$\rho(\phi_1,\phi_2\|\underset{\sim}{Y})$	$E(Y_{n+1}\|\underset{\sim}{Y})$	$V(Y_{n+1}\|\underset{\sim}{Y})$
30	0.4648	0.0397	0.4288	-0.9188	0.0473	-0.2367	0.3589	-1.1180	0.7244
50	0.5731	0.0211	0.2535	-0.7163	0.0211	-0.2028	0.0410	0.3695	0.6041
70	0.4651	0.0145	0.2589	-0.7168	0.0147	-0.1691	0.0744	-0.4170	0.5188
90	0.4637	0.0113	0.2297	-0.6397	0.0115	-0.1676	0.0670	-0.4891	0.5379

TABLE XXIV

MODERATE SAMPLE BEHAVIOR OF BOX AND JENKINS PROCEDURE FOR $\phi_1 = -.5$ AND $\phi_2 = .5$

N	$\hat{\phi}_1$	$V(\hat{\phi}_1)$	$CV(\hat{\phi}_1)$	$\hat{\phi}_2$	$V(\hat{\phi}_2)$	$CV(\hat{\phi}_2)$	$\rho(\hat{\phi}_1,\hat{\phi}_2)$	$\hat{Y}_{n+1}$	$V(\hat{Y}_{n+1})$
30	-0.4233	0.0138	-0.2775	0.5311	0.0418	0.3850	0.7180	0.4641	0.5537
50	-0.4582	0.0192	-0.3024	0.3748	0.0190	0.3677	0.7380	-0.4869	0.5857
70	-0.5510	0.0137	-0.2124	0.2889	0.0136	0.4037	0.7730	0.1674	0.5160
90	-0.5569	0.0104	-0.1831	0.3153	0.0104	0.3234	0.8060	0.0368	0.5171

TABLE XXV

MODERATE SAMPLE BEHAVIOR OF BOX AND JENKINS PROCEDURE FOR $\phi_1 = -.5$ AND $\phi_2 = 0$

N	$\hat{\phi}_1$	$V(\hat{\phi}_1)$	$CV(\hat{\phi}_1)$	$\hat{\phi}_2$	$V(\hat{\phi}_2)$	$CV(\hat{\phi}_2)$	$\rho(\hat{\phi}_1,\hat{\phi}_2)$	$\hat{Y}_{n+1}$	$V(\hat{Y}_{n+1})$
30	-0.4039	0.0396	-0.4927	-0.2754	0.0432	-0.7547	0.3880	-0.6048	0.5349
50	-0.4364	0.0215	-0.3360	-0.1253	0.0215	-1.1702	0.3830	-0.2730	0.5730
70	-0.5125	0.0145	-0.2349	-0.1698	0.0145	-0.7089	0.4370	-0.1508	0.5131
90	-0.5354	0.0113	-0.1984	-0.1374	0.0116	-0.7836	0.4650	-0.1786	0.5219

TABLE XXVI

MODERATE SAMPLE BEHAVIOR OF BOX AND JENKINS PROCEDURE FOR ϕ_1 = .3 AND ϕ_2 = .1

N	$\hat{\phi}_1$	$V(\hat{\phi}_1)$	$CV(\hat{\phi}_1)$	$\hat{\phi}_2$	$V(\hat{\phi}_2)$	$CV(\hat{\phi}_2)$	$\rho(\hat{\phi}_1,\hat{\phi}_2)$	$\hat{Y}_{n+1}$	$V(\hat{Y}_{n+1})$
30	0.3229	0.0433	0.6444	-0.2540	0.0489	-0.8706	-0.1170	0.2468	0.5720
50	0.3598	0.0216	0.4082	-0.0784	0.0216	-1.8746	-0.3270	0.2956	0.5891
70	0.2702	0.0146	0.4480	-0.1417	0.0147	-0.8556	-0.2400	0.1220	0.5085
90	0.2786	0.0115	0.3850	-0.0740	0.0115	-1.4492	-0.2500	-0.0290	0.5206

TABLE XXVII

MODERATE SAMPLE BEHAVIOR OF BOX AND JENKINS PROCEDURE FOR ϕ_1 = .5 AND ϕ_2 = -.5

N	$\hat{\phi}_1$	$V(\hat{\phi}_1)$	$CV(\hat{\phi}_1)$	$\hat{\phi}_2$	$V(\hat{\phi}_2)$	$CV(\hat{\phi}_2)$	$\rho(\hat{\phi}_1,\hat{\phi}_2)$	$\hat{Y}_{n+1}$	$V(\hat{Y}_{n+1})$
30	0.5662	0.0362	0.3360	-0.7072	0.0370	-0.2720	-0.2630	0.2396	0.5920
50	0.5090	0.0144	0.2362	-0.5859	0.0145	-0.2055	-0.3170	0.3769	0.6072
70	0.4253	0.0094	0.2280	-0.6103	0.0096	-0.1605	-0.2640	-0.3856	0.5275
90	0.4918	0.0074	0.1749	-0.6053	0.0074	-0.1421	-0.3050	-0.4079	0.5330

CHAPTER VI

SUMMARY

The main objective of this thesis is to develop a convenient theory in order to analyze the general moving average processes using the Bayesian approach. The difficulty with moving average processes is that statistical inferences about the parameters of the model must be done numerically because the posterior joint and marginal distributions of the parameters are not standard distributions. This requires a convenient and simple form for the likelihood function, and this has not been done because there is not an analytic form for the precision matrix or for the determinent of the covariance matrix.

In this study, the exact theoretical and analytical forms for the posterior distributions of the first order moving average process have been constructed. Although the marginal posterior distribution of the coefficient ϕ_1 is not standard, it has been shown that its form can be obtained in an analytically convenient form in terms of the parameter ϕ_1 directly. Also it is shown that the posterior conditional distribution of the error precision τ given the coefficient ϕ_1 is a gamma distribution with parameters

$$\frac{n + 2a}{2} \qquad \text{and} \qquad \frac{\Sigma w_j^2 \lambda_j^{-1}(\phi) + 2b}{2}$$

and that the conditional predictive density of one step ahead Y_{n+1} given

the coefficient ϕ_1 is t distribution with n + 2a degrees of freedom, $d^{-1}(\phi)e(\phi)$ location parameter, and $\frac{d(\phi)(n+2a)}{c(\phi) - e^2(\phi)d^{-1}(\phi)}$ precision, if the gamma distribution with parameters a and b is used as a prior distribution for τ. Furthermore, it was shown that the marginal expectation and variance of both the error variance τ^{-1} and the next observation Y_{n+1} can be computed.

With respect to the general moving average processes, it has been shown that it is possible to analyze these processes by setting the initial values of the errors to zero. This procedure gives an approximate covariance structure for the first q observations. Thus as the sample size increases, the recommended approach to analyze MA(q) processes is expected to be effective and reasonable. It has been shown also that statistical Bayesian inferences about the parameters can be done approximately using a t distribution in q dimensions. Based on the t approximation, a consistent Bayesian methodology for identification, estimation, and forecasting can be developed.

To demonstrate the idea of using the exact conditional likelihood function and the t approximation, two numerical studies were given to analyze the MA(1) and MA(2) processes in Chapter V. The exact procedure, the t approximation procedure, and Box and Jenkins' procedure have been used to analyze the data.

It has been shown that the numerical results obtained by using the t approximation are consistent with the theoretical objectives which supports the adequacy of the proposed approximation in analyzing the general moving average processes for moderate and large sample sizes.

It may be important, at the end of the study, to give some further research problems which are closely connected with our research. The problem of developing the exact Bayesian analysis of the ARMA(p,1) processes using the exact technique proposed in Chapter III could be studied. The exact conditional likelihood function procedure may be used to analyze ARMA(p,q) processes. Also, the problem of developing a complete Bayesian methodology to analyze ARMA(p,q) processes using the proposed t approximation can be studied.

BIBLIOGRAPHY

(1) Ali, M. M. (1977). "Analysis of autoregressive moving agerage models: estimation and prediction." Biometrika, Vol. 64, p. 535-545.

(2) Ansley, C. F. (1979). "An algorithm for the exact likelihood of mixed autoregressive moving average process." Biometrika, Vol. 66, p. 59-65.

(3) Aoki, Masanao (1967). Optimization of stochastic systems, _Topics in Discrete Time Systems_. New York: Academic Press.

(4) Box, G. E. P. and G. M. Jenkins (1970). _Time Series Analysis, Forecasting and Control_. San Francisco: Holdenday.

(5) Dent, W. T. (1977). "Computation of the exact likelihood function for an ARIMA Process." J. Statistical Computations and Simulation, 5, p. 193-206.

(6) Grenander, U. and G. Szegö (1958). _Toeplitz Forms and Their Applications_. University of California Press, Berkeley and Los Angeles.

(7) Harrison, P. J. and C. F. Stevens (1971). "A Bayesian approach to short term forecasting." Operational Research Quarterly, Vol. 22, p. 341-362.

(8) Harrison, P. J. and C. F. Stevens (1976). "Bayesian forecasting" (with discussion). Journal of the Royal Statistical Society, Series B, Vol. 38, p. 205-247.

(9) Hillmer, Steven C. and George C. Tiao (1979). "Likelihood function of stationary multiple autoregressive moving average models." Journal of the American Statistical Association, Vol. 74, No. 367, p. 652-660.

(10) Jeffreys, H. (1961). _Theory of Probability_. London: Oxford University Press.

(11) Ljung, G. M. and G. E. P. Box (1976). "Studies in the modeling of discrete time series--maximum likelihood estimation in the autoregressive moving average model." Technical Report No. 476, Department of Statistics, University of Wisconsin, Madison.

(12) Mcleod, A. I. (1977). "Improved Box-Jenkins estimators." Biometrika, Vol. 64, p. 531-534.

(13) Newbold, Paul (1973). "Bayesian estimation of Box-Jenkins transfer function-noise models." Journal of the Royal Statistical Society, Series B, Vol. 35, No. 2, p. 323-336.

(14) Newbold, Paul (1974). "The exact likelihood function for a mixed autoregressive moving average process." Biometrika, Vol. 61, p. 423-426.

(15) Nicholls, D. F. and A. D. Hall (1979). "The exact likelihood function of multivariate autoregressive moving average model." Biometrika, Vol. 66, No. 2, p. 259-264.

(16) Peterka, V. (1981). "Bayesian system identification." Automatica, Vol. 17, No. 1, p. 41-53.

(17) Phadke, M. S. and G. Kedem (1978). "Computation of the exact likelihood function of multivariate moving average models." Biometrika, Vol. 65, p. 511-519.

(18) Shore, R. W. (1980). "Bayesian analysis in economic theory and time series analysis." Pages 83-179 of <u>Bayesian Analysis in Economic Theory and Time Analysis</u>, Vol. 2 of <u>Studies in Bayesian Econometrics</u>, edited by Zellner and Kadane. Amsterdam: North Holland Publishing Company.

(19) Siddiqui, M. (1958). "On the inversion of the sample covariance matrix in a stationary autoregressive process." Ann. Math. Statist., Vol. 29, p. 585-588.

(20) Smith, J. Q. (1979). "A generalization of the Bayesian steady forecasting model." Journal of the Royal Statistical Society, Series B, Vol. 41, No. 3, pp. 375-387.

(21) Whittle, P. (1951). <u>Hypothesis Testing in Time Series Analysis</u>. (uppsala: Almqvist and Wiksells).

(22) Wilson, G. T. (1973). "The Estimation of Parameters in Multivariate Time Series Models." Journal of the Royal Statistical Society, Ser. B, 35, 76-85.

(23) Wise, J. (1955). "The autocorrelation function and the spectral density function." Biometrika, Vol. 42, p. 151-159.

(24) Zellner, A. (1971). <u>An Introduction to Bayesian Inference in Econometrics</u>. New York: John Wiley and Sons, Inc.

(25) Zellner, A. and A. Reynolds (1978). "Bayesian Analysis of ARMA Models", presented at the Sixteenth Seminar on Bayesian Inference in Econometrics, June 2-3, 1978.

VITA 2

Samir Moustafa Shaarawy

Candidate for the Degree of

Doctor of Philosophy

Thesis: A BAYESIAN ANALYSIS OF MOVING AVERAGE PROCESSES

Major Field: Statistics

Biographical:

Personal Data: Born in Cairo, Egypt, July 17, 1950, the son of Mr. Moustafa Shaarawy.

Education: Graduated from Khalil Agha High School, Cairo, Egypt, 1969; received Bachelor of Science degree in Statistics from Cairo University, Cairo, Egypt, in May, 1973; received Master of Science degree in Statistics from Oklahoma State University in May, 1981; completed requirements for the Doctor of Philosophy degree at Oklahoma State University in July, 1984.

Professional Experience: Demonstrator at Faculty of Economic and Political Science, Cairo University, Giza, Egypt, 1973-1978; Research Assistant, American University in Cairo, Egypt, 1973-1974, Research Assistant, Oklahoma State University, 1983-1984.