http://www.ejurnal.undana.ac.id/AJES ©AJES-Acad. J. Edu. Sci Ekowati /AJES, 2020, 4 (2): 22 – 25
ISSN-2654-5624(O), ISSN-2654-5969(P)

International Journal
AJES
Academic Journal of Educational Sciences
Journal Homepage: http://www.ejurnal.undana.ac.id/AJES

BAYESIAN ESTIMATION FOR ARMA MODELS

By Ch. Krisnandari Ekowati
ekowatichristine@gmail.com
Teacher Training and Educational Science Faculty, Nuca Cendana University - Indonesia

ABSTRACT: The concept of estimating a parameter is needed to help estimate a situation or observational data before making a decision. There are estimation methods that have been developed, namely the moment method, which is the oldest method, the maximum likelihood method (MLE), and the Bayes method, which is the latest method in determining the estimator of a parameter. Furthermore, the concept of forecasting is also one of the important ways to make a decision. Time series analysis technique (time series) is one of the forecasting methods that are often used, were specifically selected ARMA models. In the Bayesian approach, the parameters in the ARMA model are seen as quantities whose variance is represented by a probability distribution called a prior distribution. Within the framework of Bayes decision theory, estimator selection can be thought of as a problem of decision theory in uncertain circumstances. By using a multivariate Wishart normal prior distribution, the Bayesian estimator for ψ is ψ = z + u* and the Bayesian estimator for τ is: $\hat{\tau} = (\alpha + 1)L'S * L$, with L = (1,0,…,0)' $\in R^{n-p}$ and $S_* = S + \frac{u}{v+1}(u-y)(u-y)'$.Using the Gamma multivariate prior normal distribution, the Bayesian estimator for ψ is ψ = Z + u_o and the Bayesian estimator for τ is : $\hat{\tau} = \frac{[\alpha + \frac{n-p}{2}}{\beta}$, with u_0 = (s u + R y), $\beta_* = \beta + \frac{1}{2}(u_0 - u)'s(u_0 - u)$. Forecasting one step ahead, namely: $\hat{y}_n(I) = \sum_{i=1}^{p} \emptyset i \, y_{t-j} + \sum_{j=1}^{q} \theta j e_{t-j}$

Keywords: *Bayesian estimation, forecasting, time series, ARMA model.*

INTRODUCTION

Forecasting is one of the essential elements in decision making because the effectiveness of a decision depends on several factors that cannot see when the decision taken. The role of forecasting explores into many fields such as economics, finance, state administration, geophysics and population (Abraham, 1983).

The time series analysis technique which is one of the forecasting methods is used in many fields of scientific work, especially in making operational forecasting (Box & jenkins, 1970).

The ARMA (*Antoreg Restive Moving Average*) model has been used widely as a statistical model. The statistical model is a mathematical model that contains heretical terms to accommodate the inevitable differences. This heresy interpreted as a random variable that has a probability distribution with a zero mean. The problem is how to estimate the parameters of the ARMA model and use the model to calculate the forecast value for the future. In this case, there are two approaches, namely the classical approach and the Bayesian approach (Box & Tiad, 1973).

In the classical approach, the parameter in the ARMA model is an unknown fixed amount. Random samples are taken from the population and based on observed prices of the samples; knowledge of parameters in the ARMA model obtained (Abraham Bovas, 1983).

In the Bayesian approach, the parameters in the ARMA model seen as quantities whose variance is represented by a probability distribution called a prior distribution. It is a subjective distribution, based on one's beliefs and formulated before data is taken. Then the sample is drawn from the population, and the prior distribution adjusted to the information of this sample, the adjusted priors are called the posterior distribution. This adjustment is done using Bayes rules (Box & Tiad, 1973).

The two approaches above are informal because they do not produce the best single estimator. The selection of estimators, according to the classical method, is a matter of whether it cannot be more important than efficiency and so forth. The same problem is also faced by Bayesians to choose estimators such as posterior mean, mode or median. If what desired is a general idea of price parameters, then an informal approach is sufficient. However, ifit faces a situation where estimation errors will have serious consequences, then a formal approach is needed. If the loss function can be determined then the point estimator can be done according to the decision theory far obtained (Lehman, 1983).

Based on the background of the problem outlined above, the problem that will discuss in this paper is how to estimate parameters in the ARMA time series model through the Bayesian approach.

The purpose of writing this scientific work is;

1. To understand how the ARMA time series model;
2. To know that to estimate the parameters of the ARMA model in addition to the *likelihood* function approach can also use the Bayesian approach.

The benefits of writing this scientific work are:

1. By knowing the problems in the ARMA time series model, is expected to provide an alternative to solving the estimation problem in the ARMA time series model.
2. Can provide input to readers, especially those who are relevant to this paper.

DISCUSSION

The writing method of this scientific paper is a study of literature.

1. Determine Bayesian Estimator

Within the framework of Bayes decision theory, estimator selection can think of as a problem of decision theory in uncertain circumstances. Here the decision corresponds with the possible estimators. The loss function is determined to provide losses for errors in estimating estimated parameters. If it is possible to determine the loss function, then the best action is the action that minimizes the expected loss function (Lehman E.L, 1983).

Theorem 1.1 For the quadratic loss function, if E (θ) there is a Bayes estimator for θ is E (θ) (Box & Tiad, 1973).

Proof:

Because k > 0, to minimize R$(\theta - \delta)$ it is sufficient to minimize E$(\theta - \delta)^2$.

Because $-\theta = \delta - E(\theta) + E(\theta) - \theta$, then:

$$E(\theta - \delta)^2 = E(\theta - E(\theta) + E(\theta) - \delta)^2$$

$$= E[\delta - E(\theta)]^2 + 2E[(\delta - E(\theta))(E(\theta) - \theta)] + E[E(\theta) - \theta)^2$$

$$=[\delta - E(\theta)]^2 + E[E(\theta) - \theta)^2$$

Since the last term does not meet δ, the minimum price depends only on the term $[\delta - E(\theta)]^2$. The term will be zero only if $\delta = E(\theta)$. So this price is the price that minimizes the risk function for each θ.

Furthermore, if the parameter is a vector, that is $\theta = (\theta_1, \theta_2, \dots, \theta_k)^1$ with k ≥ z. In this case, the loss function d εR^k and the random variable $X_1, \dots, X_k \varepsilon R^2$. Bayes estimator for θ will be found in the following theorem.

Theorem1.2 For the quadratic loss function, if E(θ) and Kov (θ) exist, the Bayes estimator θ is E(θ.) (Box & Tiad, 1973).

2. ARMA Time Series Model

Consider the time series model (Abraham, 1983):

$$y_t = \sum_{i=1}^{p} \emptyset i\, yt - i + \sum_{i=1}^{q} \theta j^e t - j + e_t$$

Which $\{e_t\}$ is the normal random i.i.d random variable with $e_t\, \tau^{-1}$ with $\tau > 0$ and unknown. We write the model (2.1) with ARMA (p, q).

Suppose there are n observations $(y_1, y_2, \dots, y_n)^1$, then the residue is written as:

$$e = y_t - \sum_{i=1}^{p} \emptyset i\, yt - i + \sum_{i=1}^{q} \theta j^e t - j + e_t \qquad (2.2)$$

With the condition that the first p observation and suppose $e_p = e_{p-1} = \dots = e_r = 0$, with r = min (0, p+1 – q).

Then, the complete model can be written in the form of a matrix as follows:

$$e = y - z\psi \qquad (2.3)$$

Where:

$e = (e_{p+1}, e_{p+2}, \dots, e_n)'$

$y = (y_{p+1}, y_{p+2}, \dots, y_n)'$

$\psi = (\emptyset_1, \emptyset_2, \dots, \emptyset_p,\ \theta_1, \theta_2, \dots, \theta_q)'$ dan

$$z = \begin{pmatrix} y_p & y_{p-1} & \cdots y_1 & e_p & e_{p-1} & \cdots & y_{p+1-q} \\ y_{p+1} & y_p & \cdots y_2 & e_{p+1} & e_p & \cdots & y_{p+2-1} \\ & & & \vdots & & & \\ y_{n-1} & y_{n-2} & \cdots y_{n-p} & e_{n-1} & e_{n-2} \cdots & e_{n-q} & \end{pmatrix}$$

With residues assumed to have the following properties:

(1) $E(e_j) = 0$
(2) Var $(e_j) = \tau^{-1}$ dan
(3) Kov $(e_j . e_k) = 0$, j ≠ k

3. Prior and Posterior Analysis

In Bayesian analysis, prior distribution decisions usually play an essential role. Using the Bayes theorem, initial information expressed by prior distributions and sample information expressed by the *likelihood* function are combined to form a posterior distribution. One way to avoid this difficulty is to limit the distribution of priors in certain distributionfamilies that depend on the form of the *likelihood* function. The Bayesian statement has developed the concept of prior peer distribution. The prior normal Wishart or normal gamma distribution is usedto obtain Bayesian estimators (Dudewiez & Mishra, 1988).

Theorem 3.1(Box & Tiad, 1983):

If V> 0, m ∈ R^k, $u \in R^k$ and R matrix type k x k, then:

V (m-u)’ R(m-u) + (m-x)’ R(m-x)

$= (v+1)(m-u^*)' R(m-u^*) + fr\,[\frac{V}{V+1}(u-x)(u-x)' R]$

With $u^* = (Vu + x) / (v+1)$

Proof:

V (m-u)’ R(m-u) + (m-x)’ R(m-x)

$$= \frac{1}{v+1}[v(m-u)]' R[V(m-u)] + \frac{v}{v+1}(m-u)' R(m-u)$$

$$+ \frac{1}{v+1}(m-x)' R[v(m-u)] + \frac{v}{v+1}(m-u)' R(m-x)'$$

$$= (v+1)(m-u^*)' R(m-u^*) + fr\,[\frac{V}{V+1}(u-x)(u-x)' R] \qquad (2.1)$$

Theorem 3.2 (Box & Tiad, 1983):

x a random sample of a multivariate normal distribution population with a mean vector m value is unknown and the precision matrix R is unknown. The prior distribution with m and R is $\varepsilon(m, R) = \varepsilon_1(m \mid R)\, \varepsilon_2(R)$. With $\varepsilon_1 \sim N(u, vR) \ni u \varepsilon R^k$ and u > 0 and

$\varepsilon_2 \sim W(\infty, S) \ni \infty > k-1$ of the positive definite symmetry matrix.

Then the posterior distribution with m and R is; $\Pi(m,R) = \Pi_1(m \mid R)\, \Pi_2(R)$. With $\Pi \sim n(U*, (v+1)R)$ dan $\Pi \sim W(\infty+1, S*)$ where:

u* = (u +x)/(v+1)

S* = S + $\frac{v}{v+1}(u-x)(u-x)'$

Theorem 3.3 (Box & Tiad, 1983):

Given the *likelihood* function f(x|m,R) $\infty\, [\det(R)]\frac{1}{2}\exp\left[-\frac{1}{2}(x-m)'R(x-m)\right.$ and the posterior density function $\varepsilon(m,R \propto [\det(R)]^{\frac{1}{2}}\exp\left[-\frac{1}{2}v(m-u)'''R(m-u)\right].[\det(R)]^{(\alpha-K-1)/2}\exp -\frac{1}{2}he\,(SR)]$

then the marginal posterior distribution R is the Wishart distribution with degrees of freedom $\propto +1$ and a precision matrix

S* = S + $\frac{v}{v+1}(u-x)(u-x)'$

Proof:

Π(R)

$\propto \int_{n\varepsilon R^k}\{[\det(R)]^{\frac{1}{2}}\exp\left[\frac{1}{2}(v+1)(m-u*)'R(m-u*)\right][\det(R)]^{\frac{\alpha-k}{2}}\exp[-\frac{1}{2}fr(S*R]\}$

dm.

Because:

$\int_{n\varepsilon R^k}\{[\det(v+1)R]^{\frac{1}{2}}\exp\left[\frac{1}{2}(v+1)(m-u*)'R(m-u*)\right]dm\} = 1$

Then

$\Pi(R) \propto [\det(R)]^{\frac{\alpha-k}{2}}\exp[-\frac{1}{2}fr(S*R]\}$ (2.4)

Relation (2.4) is a Wishart distribution with degrees of freedom $\propto +1$ and a precision matrix S *

As a result of 2.6:

Given the *likelihood* function:

$L(\psi,\tau|s_n) \propto \tau^{\frac{(n-p)}{2}}\exp[-\frac{1}{2}\tau(Y-Z\psi)'(Y-Z\psi)$ (2.5)

and the density function prior $\varepsilon(m,R \propto [\det(R)]^{\frac{1}{2}}\exp\left[-\frac{1}{2}v(m-u)\;\; R(m-u)\right][\det(R)]^{\frac{(\alpha-k-1)}{2}}\exp[-\frac{1}{2}fr(SR]\}$ (2.6)

then the Bayes estimator for τ is

$\hat{\tau} = (\alpha+1)L'S*L,$ dengan L = (1,0,…,0)' $\in R^{n-p}$ and

S* = S + $\frac{u}{v+1}(u-y)(u-y)'$

Proof:

Use theorem 2.5 by taking R= τI, m=Zψ dan k= n-p, then the posterior marginal distribution τI is:

$\Pi(\tau I) \propto [\det(\tau I)]^{\frac{(\alpha-n+p)}{2}}\exp[-\frac{1}{2}fr(S.\tau I]$

According to the theorem of the Wishart distribution, namely: if $V \sim W_k(T, n)$ and L are a vector, then L' V L ~ L' $T^{-1}L\hbar^2_{(n)}$ with L = (1,0,…,0)'.

Member of R^{n-p}, the Bayes estimator for τ is $\hat{\tau} = E(\tau)$ or $\hat{\tau} = (\alpha+1)L'S*L$

As a result of 2.7:

Given the likelihood function (2.5) and the prior density function (2.6), the Bayes estimator for ψ is ψ = $Z^+ u_*$ with $u_* = (vu + y)/(v+1)$

Proof:

Taking R = τI and using theorem 2.6, m = Zψ and k= n-p, the marginal distribution of posterior Zψ is:

Π(Z ψ)

$\propto [1+(v+1)(Z\psi-u*)'^{S*^{-1}}(z\psi-u*)]^{\frac{-(\alpha+2)}{2}}$

With $S_* = S + \frac{u}{v+1}(u-y)(u-y)'$

According to theorem 2.2, the Bayes estimator for Z ψ adalah $z\widehat{\psi} = E(Z\psi)$. According to the theorem about the expectations of a location vector, then $z\widehat{\psi} = u_*$.

The best approach to the equation A x = y is xo = A y, so the Bayes estimator approach for ψ is ψ = $z^+ u_*$.

Theorem 3.4 (Box & Tiad, 1983):

x is a random sample of a multivariate normally distributed population with an unknown vector m mean and a precision matrix W R, with a certain positive definite symmetry matrix R and W unknown.

The joint distributions of prior m and w are:

$\varepsilon(m,w) = \varepsilon_1(m\,|w)\,\varepsilon_2(w)$ where $\varepsilon_1$is a multivariate normal distribution with mean vector u and precision matrix W S $\ni u \in R^k$and S is a positive definite symmetry matrix of type kxk and $\varepsilon_2$is gamma distribution with parameters α dan $\beta \ni \alpha > 0 \text{ and } \beta > 0$.

Then the joint distribution of posterior m and w given x is:

$\Pi(m,w,x) = \Pi_1(m \mid w)\,\Pi_2(w)$where $\Pi_1$is a multivariate normal distribution with mean vector u^c and a precision matrix w (S + R),

where$u^c = (S+R)^{-1}(S\,U + R\,X)$ and $\Pi_2$is the gamma distribution with parameters $\alpha + k/2$ and $\beta*$, with $\beta = \beta + 1/2(u^0-u)^1S(u^0-u)$.

Theorem 3.5 (Box & Tiad, 1983):

Given the likelihood function: f(m,w|x) $\alpha[\det(WR)]^{\frac{1}{2}}\exp[-\frac{1}{2}(x-m)'W\,R(x-m)]$and the prior density function:

$\varepsilon(m,w)\,\alpha[\det(w\,s)]^{\frac{1}{2}}\exp\left[-\frac{1}{2}(m-u)'w\,s(m-u)\right].w^{(\alpha-1)}\exp$

Then the marginal posterior vector mean m distribution is a multivariate t distribution with degrees of freedom $2(\alpha + k/2)$, location vector u_0,

where$u_0 = (S + R)^{-1}(s\,u + R\,x)$

Effect 2.10:

Given the *likelihood* function: L $(\psi,\tau|s_n)$ $\propto \tau^{\frac{(n-p)}{2}}\exp[-\frac{1}{2}\tau(Y-z-\psi)'(Y-z-\psi)$and the prior density function on theorem 2.9, the Bayesian estimate for ψ is $\psi = z^+ u_0$ with $u_0 = (S + R)^{-1}(s\,u + R\,y)$

Proof:

Using theorem 2.9 and taking R=I, w = τ, Z=ψ dan k = n-p, then the marginal distribution of posterior z $\psi$is a multivariate t distribution with degrees of freedom $2(\alpha + (n-p)/2)$, location vector u0 and precision matrix

$$\{[\alpha + (n-p)/2](s + R)\}$$

According to theorem 2.2, the Bayes estimator for z ψ = E (z ψ) or z ψ =u_0. So the Bayes approach for ψ adalah $\hat{\psi} = z^+ u_0$.

4. ARMA Process Forecasting

Estimation models that have assessed for time series data they have, will be used for time series inference in the future based on their past time behaviour. Based on a model, we want to reduce the conditional distribution of future observations if known past observations. This step, which is the last step in forming a time series model, is known as forecasting. If the best forecast is intended as a forecast that minimizes the forecast of the heresy squared mean, then the conditional distribution of observations that will come with the mean distribution is a good forecast (Box & Jenkins, 1970).

We write $\hat{y}_n$(I) the expected value of $y_n + 1$, if it is known that the previous time series has reached time n. the forecast is one step ahead, namely:

$$\hat{y}_n(I) = E(y_n + 1|s_n) = \sum_{i=1}^{p} \emptyset i \; y_{t-j} + \sum_{j=1}^{q} \theta j e_{t-j}$$

Then the forecast error is $[e_n(1)|s_n] = y_{n+1} - \hat{y}_n(I) = e_{n+1}$ and the forecast error variance is

$\text{Var}[e_n(1)|S_n] = \tau^{-1}$.

As a result, the size of disperse depends only on the parametersψ *and* τ. Therefore by using the estimated values, the confidence interval for future observations can be calculated. Because $e_t$is normally distributed, we can describe the entire distribution of future heresies and thus future observations. In this case;

$$[e_n(1)|S_n] \sim N(o, \tau^{-1}) \text{ atau } [y_{n+1}|S_n] \sim N(\hat{y}_n(I), \tau^{-1}).$$

Therefore, a 100 (1-2d)% confidence interval forecast can be written as:

$$\{\hat{y}_n(I) - Z_d SD[e_n(1)|S_n], \hat{y}_n(I) - z_d SD[e_n(1)|S_n]\}$$

Where fn (z_d) = 1- d and fnare the standard normal random variable cumulative distribution functions.

CONCLUSIONS

Based on the discussion of the problem in Chapter II, it can be concluded that:

1. Using the Wish normal multivariate Wishart distribution, the Bayesian estimator for ψisψ = z + u* and the Bayesian estimator for τadalah :$\hat{\tau} = (\alpha+1)L'S * L$, dengan L = (1,0,…,0)' $\in R^{n-p}$and

   $$S_* = S + \frac{u}{v+1}(u-y)(u-y)'$$
2. Using the Gamma multivariate prior normal distribution, the Bayesian estimator for ψisψ = Z + $u_o$and the Bayesian estimator for τ is :$\hat{\tau} = \frac{[\alpha+(n-p)/2}{\beta}$, with $u_0 = (s\,u + R\,y)$,

   $$\beta_* = \beta + \frac{1}{2}(u_0 - u)'s(u_0 - u)$$
3. Forecasting one step ahead, namely:

   $$\hat{y}_n(I) = \sum_{i=1}^{p} \emptyset i \; y_{t-j} + \sum_{j=1}^{q} \theta j e_{t-j}$$

SUGGESTIONS

The Bayesian approach can use to determine the predictive distribution for y_{n+1}. A conditional distribution is required given past observations and also a posterior distribution for the parameters of the ARMA model (p, q). This is an interesting theme for readers, especially fans of statistics, to continue this paper.

REFERENCES

Abraham Brovas & Ledolter Johannes, (1983), *Statistical Methods for Forcasting*, john Wiley & Sons,Inc.

Edward, J. Dudwiez&Satya, N, Mishra. (1988). *Modern Mathematics Statistics*.John Wiley. Canada

George E.P. B ox & George C. Tiad. (1973). *Bayesian Inference In Statistical Analysis*. Addison-Wesley. London

George E.P. Box &Guilym M. Jenkins.(1970). *Time Series Analysis (Forecasting and Control)*.Holden-Day. California

Lehmann E.L, (1983), *Theory of Point Estimation*, John Wiley & Sons, Inc.

Wald Abraham, (1966), *Sequential Analysis*, John Wiley & Sons,Inc