# اشـتقــاق دالـة تنبــؤ بيــزية لنمـاذج
# ARMAX الموسمــية

دكتور / جمال أحمد محمد الشوادفى
كلية التجارة بنين - جامعة الازهر
قسم الاحصاء والرياضيات

**إذا** كانت لديك سلسلتين زمنيتين $(y_t , x_t)$ تعتمد احداهما ( $y_t$ )على الاخرى( $x_t$ ) من خلال نموذج ARMAX الموسمى وكنت بحاجة الى التنبـؤ بـالقيم المستقبلية للسلسلة $y_t$ فإن هذا البحث يقدم طريقة - باستخدام اسلوب بيز - للحصول على مثل هـذه التنبـؤات . وتتميز هذه الطريقة بأنها تجمع بين ميزة استخدام اسلوب بيز ، واحد أهم النمـاذج المناسبة لتحليل السلاسل الزمنية فى مجالات الاقتصاد والاعمال والبيئة وغيرها ، وهـو نموذج ARMAX الموسمى . كما أن استخدام طريقة مناسبة للتنبـؤ سيكون مفيدا جدا لصانعى القرارات فى مثل هذه المجالات .

**وتتمثل** المشكلة الرئيسية فى هذا البحث فى صعوبة التعامل الرياضى والعددى مع دالة الامكان فى هذه الحالة ، وصعوبة اختيار دالة توزيع قبلية مناسبة لهـا . وينتـج عن ذلك عدم وجود صيغة قياسيـة لـدوال الاحتمـال التنبؤيـة . وقد أمكـن التغلب علـى هذه المشـكلة باستخدام دالة مناسبة كتقريب لدالة الامكان الشرطية ثم اختيار التوزيع القبلى الملائم لها . والتوزيـع القبلـى يـأخذ صـورة دالـة توزيـع جامـا-المعتـدلNormal - Gamma density أو صــورة دالــة عديمــة المعلومــات Non-Informative prior density.

**وبنـاءا** على دالة الامكان المقترحة والتوزيع القبلى المرافق ، فقد أمكـن اثبـات أن دالـة التنبـؤ الحدية للمشـاهدة المستقبلية الاولـى تـأخذ شكل توزيـع t غـير المركـزى Non-central univariate t distribution ، كذلك فانه أمكن اثبـات أن دالـة التنبـؤ الشرطية للمشاهدات المستقبلية بعد الاولى سـتأخذ أيضا شـكل توزيـع t غير المركـزى ، ومن ثم يمكن بسهولة استخدام الطريقة المقترحة لايجاد التوقع وانشـاء فـترات ثقـة لهـا أعلى كثافة تنبؤية HPD regions لأى مشاهدة مستقبلية.

(11) Lutkepohi , Helmut (1993)
Introdution to Multiple Time Series Analysis . Springer Verlag , Berlin , Heidlberg ,Germany.

( 12) McGraw , G. A. and et al (1993)
Conditions for the Equivalence of ARMAX and ARX Systems , IEEE Transactions in Automatic Control , Vol. 38 , No. 4 , pp. 632- 636

(13) Mills ,Ternce C. (1991)
Time Series Techniques for Economists , Cambridge Univerisity Press Cambridge , England.

(14) Nerlove , M. (1972)
"lags in Economic Behaviour " Econometrica 39 , pp. 359-382.

(15) Newbold , Paul (1974)
The Exact Likelihood function for a Mixed Autoregressive Moving Average Process , Biometrika , Vol. 61, pp. 423-426.

(16) Spanoss, A.(1986)
Statistical Foundation of Econometeric Modeling ,Cambridge University Press , Cambridge ,U.K.

(17) Wei ,William W.S. (1990)
Time Series Analysis Univariate and Multivariate Methods , Addison Wisely Publishing Company , Inc. Redwood City , California , U.S.A.

(5) Chatfield (1980)
The Analysis of Time Series , Theory and Practice , Chapman and Hill Ltd. , London.

(6) Deistler , M. (1980)
The Properties of the Parameterization of ARMAX Systems and Their Relivance for Structural Estimation and Dynamic Specification , The Fourth World Congress of the Econometeric Socitey in Aix-en - Provence , France.

(7) Gaynor ,Patricia E. and Kirkpatrick Rickey C.(1994)
Time Series Modeling and Forecasting in Business and Economics, McGrawHill , Inc. New York , U.S.A.

(8) Harvey A.C. (1981)
The Econometeric Analysis of Time sereis , Phlip Allan Puplishers Limited ,Market Place ,Deddincton Oxford Ox5 45E, Great Britian.

(9) Hendry , D.F. and et al (1985)
" Dynamic Specification " in Z. Grilichess and M.S. Intriligator (Editors), Handbook of Econometrics ,volume 11 pp.1023 -1100 , Amestrdam North-Holand.

(10) Judge G.G. and et al (1985)
The Theory and Practice of Econometrics , 2nd Edition ,Wiley New York , U.S.A.

## 4- CONCLUSIONS:

The main purpose of this paper is to develop a convenient Bayesian approach to forecast the future observations for the seasonal ARMAX model. This was achieved by approximating the conditional liklihood function by a normal gamma function . Thus the marginal predictive distribution of the first future observation is approximated by a non-central univariate t distribution , and the conditional predictive distribution of observations at higher lags are approximated by a non-central univariate t distribution. In addition it was shown that the highest predictive density region (HPD region) for any future observation may be constructed.

## REFRENCES


(1) Ali M. M. (1977)
Analysis of Autoregressive Moving Average Models: Estimation and Prediction , Biometrika , Vol. 64 , pp. 535-545.

(2) Bierens , Herman , J. (1987)
ARMAX Model Specification Testing With an Application to Unemployment in the Netherlands, El-Savier Science Publishers B.V. North Holland.

(3) Broemeling , L. D. and Shaarawy , S. (1985)
A Bayesian Analysis of Time Series , Contribution to Economic Analysis Studies in Bayesian Econometrics and Statistics in Honor of Bruno de Finetti , Edited by Zellner and Goal , North Holland , Amsterdam.

(4) Box, G. E. P. and Jenkins , G. M. (1976)
Time Series Analysis , Forecasting and Control ,Holden Day , San Francisco.

$$E(y_{n+2} \mid y_{n+1}, S_n) + t_{(\alpha/2,\, n+2a-m+1)} \cdot [\, p(y_{n+2} \mid y_{n+1}, s_n)]^{-0.5}$$

...... (3.8)

where $E(y_{n+2} \mid y_{n+1}, S_n)$ and $p(y_{n+2} \mid y_{n+1}, s_n)$ are as it defined in (3.5) and (3.6) respectively, $0 < \alpha < 1$, and t is obtained from the student's tables.

Thus, using the condiatinoal predictive density of $y_{n+k}$, $k > 2$, given $y_{n+1} = y_{n+1}$, $y_{n+2} = y_{n+2}$, ........, $y_{n+k-1} = y_{n+k-1}$, one may forecast the kth future observation, and hence by a sequence of conditional expectations, one may forecast k steps ahead.

Suppose that one can't or unwilling to specify the parameters a, b, $\mu$, and R of the normal gamma prior density, (2.8), one might prefer to use a non-informative prior density, (2.9). For such case, the following corollary may be used.

Corollary (3.2)

If the aproximate likelihood function (2.7) is combined with the non-informative prior density (2.9), the marginal predictive density of the frist future observation, $y_{n+1}$, will be a non-central univariarte t distribution with df = n-m-K degrees of freedom, location

$$E(y_{n+1} \mid Sn) = (1 - B_1' A^{-1} B_1)^{-1} \; B_1' A^{-1} B$$

...........(3.9)

and precision :

$$P(y_{n+1} \mid Sn) = \frac{(n+K-m)\;(1 - B_1' A^{-1} B_1)}{Y'Y - B'A^{-1}B - E^2/D}$$

........ (3.10)

where A, B1 and B are as defined in (3.3).

By a sequence of conditional expectations, one may forecast k steps ahead with the aid of corollaries (3.1) and (3.2).

where $E(y_{n+1} \mid S_n)$ and $p(y_{n+1} \mid s_n)$ are as it defined in (3.1) and (3.2) respectively , $0 < \alpha < 1$ , and t is obtained from the student t tables.

The following corollary shows how one can forecast k step ahead , $k > 2$ , assuming we have n observations generated by a seasonal ARMAX model.

## Corollary (3.1)

If the approximate conditional likelihood function (2.7) is combined with the normal gamma prior density (2.8) , The conditional predictive density of the second future observation , $y_{n+2}$ , will be a non-central univariate t distribution with df = n+2a-m+1 degrees of freedom , location ;

$$E(y_{n+2} \mid y_{n+1}, Sn) = E / D \qquad \text{...............(3.5)}$$

and precision

$$P(y_{n+2} \mid y_{n+1}, Sn) = \frac{df\, D}{F - E^2/D} \qquad \text{......(3.6)}$$

where;

$S_n$ , $\mu$ , a , b and R are as it were defined in (2.7) and (2.8) , $m = \max(p+Ps, h+HS)$ , and Y, E , D , F are as it were defined in (3.3), but the quantities A,B1 , B are modified by letting n=n+1 ,$y_{n+1} = E(y_{n+1} \mid s_n)$ ,(3.2) , and

$$\hat{e}_{n+1} = y_{n+1} - Y_1' . \beta_1 + X_1' . \beta_2 + \hat{E}_1' . \beta_3 \qquad \text{.....(3.7)}$$

In addition , a highest predictive density region (HPD region ) for $y_{n+2}$ of content $(1-\alpha)$ is:

THEORM(3.1)

If the approximate conditional likelihood function (2.7) is combined with the normal gamma prior density (2.8) , The marginal predictive density of the first future observation , $y_{n+1}$ , will be a non-central univariate t distribution with df = n+2a-m degrees of freedom , location ;

$$E(y_{n+1} \mid Sn) = E/D \qquad ................(3.1)$$

and precision

$$P(y_{n+1} \mid Sn) = \frac{df\ D}{F - E^2/D} \qquad ........(3.2)$$

where

$$E = B_1'\,(A+R)^{-1}\,(B+R\mu)$$

$$D = 1 - B_1'(A+R)^{-1}B_1$$

$$F = \mu'R\mu + 2b + Y'Y - (B+R\mu)'(A+R)^{-1}(B+R\mu) \qquad ......(3.3)$$

and $S_n$ , $\mu$ , a , b and R are as it were defined in (2.7) and (2.8) , m= max(p+Ps,h+HS).

$Y = (y_{m+1}\ \ y_{m+2}\ ............\ y_n)'$

$B = A_0'\,Y$ , $A = A_0'\,A_0 + B_1B_1'$ ,and $A_0$ is (n-m).k matrix in the form $A_0 = (Y_{t-1}' : X_t' : \hat{E}_{t-1}')$ , and $B_1 = (Y_1' : X_1' : \hat{E}_1')'$ is a vector of order k.1

$Y_1 = (\ y_n\ ..............\ y_{n-p+1}\ \ y_{n-s+1} \cdots y_{n-Ps+1}\ \ y_{n-s} \cdots\cdots y_{n-p-Ps+1}\ )'$

$X_1 = (\ x_{n+1}\ x_n\ \ldots\ x_{n-h+1}\ \ x_{n-s+1} \cdots\cdots x_{n-Ps+1}\ \ x_{n-s} \cdots\cdots x_{n-h-Hs+1}\ )'$

$\hat{E}_1 = (\ \hat{e}_n\ .............\ \hat{e}_{n-h+1}\ \ \hat{e}_{n-s+1} \cdots\cdots \hat{e}_{n-Hs+1}\ \ \hat{e}_{n-s} \cdots\cdots \hat{e}_{n-q-Qs+1}\ )'$

In addition , a highest predictive density region (HPD region ) for $y_{n+1}$ of content $(1-\alpha)$ is:

$$E(y_{n+1} \mid S_n) + t_{(\alpha/2\,,\,n+2a-m)} \cdot [\,p(y_{n+1} \mid s_n)]^{-0.5} \qquad .........(3\ .\ 4)$$

where $\beta_1$ , $\beta_2$ , $\beta_3$ are as it were defined in (2.3) and B, $S_n$ , Y , $Y_{t-1}$ , $X_t$ , $E_{t-1}$ , are as it were defined in (2.5) , $\tau>0$ , $m= max(p+Ps,h+Hs)$ .

An aproparíate choice of prior density function is normal -gamma in the form;

$$H(B, \tau) \ \alpha \ \tau^{0.5(2a+k)-1} \exp(-\tau/2)((B-\mu)'R(B-\mu) +2b) \qquad .........(2.8)$$

where $\tau>0$ , $a>0$ $b>0$, $\mu$ is the vector of condtional expectation of B, given $\tau$ and R is a positsive definite matrix of order $K=p+P+pP+h+H+hH+q+Q+qQ+1$

However, if there is no prior information one may use a non-informative prior density function in the form :

$$H(B, \tau) \ \alpha \ \tau^{-1} \ , \tau>0 \qquad .........(2.9)$$

Equation (2.9) may be obtained from the normal gamma density , (2.8) , by letting $a \to -k/2$ , $b \to 0$ , and $R=0_{K.K}$

## 3. BAYESIAN FORECASTING FOR THE SEASONAL ARMAX MODEL

The Bayesian approach requires information about the parameters in the form of the proper or improper prior density function , like functions (2.8) and (2.9) respectively . The prior density function is combined with the likelihood function to give predictive inferences . The marginal predictive density of the first future observation , $y_{n+1}$ , will be introduced through the following theorm:

Many invesigations for simpliring the likelihood function can be found in Newbold (1974), Ali (1977), Broemeling and Shaarawy (1985) and others. The proposed approximation here is similar to the one used by Broemeling and Shaarawy (1985) with nonseasonal ARMA model and is based on : equating the initial values of the errors $E_0 = (e_m \; e_{m-1} \; \ldots\ldots\ldots\ldots \; e_{m-q-Qs+1})'$ by their unconditional expectation namely zero and replacing the exact residuals by their least squares estimates. The least squares estimates, say $\hat{E} = (\hat{e}_{m+1} \; \hat{e}_{m+2} \; \ldots\ldots\ldots \; \hat{e}_n)'$ are obtained by searching over the parameters space for the value of vector $B$, say $\hat{B}$, which minimizes the residuals sum of squares.

$$E'E = \sum_{t=m+1}^{n} e_t^2 \qquad \ldots\ldots (2.6)$$

where $m = \max(p+Ps, h+Hs)$.

Using the estimated residuals
$\hat{E} = (\hat{e}_{m+1} \; \hat{e}_{m+2} \; \ldots\ldots\ldots \; \hat{e}_n)'$, and equating the initial values of the errors $E_0 = (e_m \; e_{m-1} \; \ldots\ldots\ldots\ldots \; e_{m-q-Qs+1})'$ by their unconditional expectation namely zero, we can rewrite the likelihood function (2.5) approximately as ;

$$\hat{L}(B,\tau|S_n) \alpha \, \tau^{(n-m)/2} \exp(-\tau/_2)(Y-Y_{t-1}\beta_1-X_t.\beta_2-\hat{E}_{t-1}.\beta_3)'(Y-Y_{t-1}.\beta_1-X_t.\beta_2-\hat{E}_{t-1}.\beta_3) \qquad \ldots\ldots\ldots(2.7)$$

$L(B,\tau|S_n)\alpha\,\tau^{(n-m)/2}$

$$\exp(-\tau/_2)(Y-Y_{t-1}.\beta_1-X_t.\beta_2-E_{t-1}.\beta_3)'(Y-Y_{t-1}.\beta_1-X_t.\beta_2-E_{t-1}.\beta_3) \qquad \text{............ (2.5)}$$

where

$B = (\beta_1' : \beta_2' : \beta_3')'$

is a K.1 vector of parameters ,$\tau > 0$ ,

$$S_n = \left[\begin{array}{c|c} Y_0 & X_0 \\ \hline Y & X \end{array}\right]$$

K=p+P+pP+h+H+hH+q+Q+qQ+1

$Y = ( y_{m+1} \quad y_{m+2} \quad \ldots\ldots\ldots \quad y_n )'$ is the vector of n-m observations on y,

$X = ( x_{m+1} \quad x_{m+2} \quad \ldots\ldots\ldots \quad x_n )'$ is the vector of n-m observations on x,

m = max ( p+Ps , h+Hs )

The likelihood function (2.5) is conditioned on the starting values

$Y_0 = (Y_m \; Y_{m-1} \ldots\ldots\ldots\ldots \; Y_{m-p-Ps+1})'$

$X_0 = (X_m \; X_{m-1} \ldots\ldots\ldots\ldots \; X_{m-h-Hs+1})'$ , and

$E_0 = (e_m \; e_{m-1} \ldots\ldots\ldots\ldots \; e_{m-q-Qs+1})'$

However the conditional likelihood function (2.5) is analytically intractable because there is no closed form for the precision matrix or for the determinant of the covariance matrix in terms of the parameters directly . Further more , with any prior distribution the predictive densities are not standard which mean that inferences about the future observations must be done numerically. For the above reasons a simple analytic form for the the likelihood function is needed .

$$X_t = \begin{bmatrix} X_{m+1} & \cdots\cdots\cdots & X_n \\ X_m & \cdots\cdots\cdots & X_{n-1} \\ & \cdots\cdots\cdots & \\ X_{m-h+1} & \cdots\cdots\cdots & X_{n-h} \\ X_{m-s+1} & \cdots\cdots\cdots & X_{n-s} \\ & \cdots\cdots\cdots & \\ X_{m-Hs+1} & \cdots\cdots\cdots & X_{n-Hs} \\ X_{m-s} & \cdots\cdots\cdots & X_{n-s-1} \\ & \cdots\cdots\cdots & \\ X_{m-h-Hs+1} & \cdots\cdots\cdots & X_{n-h-Hs} \end{bmatrix}$$

$$E_{t-1} = \begin{bmatrix} e_m & \cdots\cdots\cdots & e_{n-1} \\ & \cdots\cdots\cdots & \\ e_{m-q+1} & \cdots\cdots\cdots & e_{n-q} \\ e_{m-s+1} & \cdots\cdots\cdots & e_{n-s} \\ & \cdots\cdots\cdots & \\ e_{m-Qs+1} & \cdots\cdots\cdots & e_{n-Qs} \\ e_{m-s} & \cdots\cdots\cdots & e_{n-s-1} \\ & \cdots\cdots\cdots & \\ e_{m-q-Qs+1} & \cdots\cdots\cdots & e_{n-q-Qs} \end{bmatrix}$$

..................(2.4)

Assuming that the errors ( $E_t$ ) are normal random variables with mean vector zero and precision matrix $I\tau$ , $I$ is an identity matrix , $\tau > 0$ , we can write the conditional likelihood function as;

The class of models (2.1) may be written in a matrix notation as:

$$Y = Y_{t-1}.\beta_1 + X_t.\beta_2 + E_{t-1}.\beta_3 + E_t \qquad \text{.................... (2.3)}$$

where :

$Y = (\, y_{m+1} \quad y_{m+2} \; \ldots\ldots\ldots\ldots \; y_n \,)'$

$E_t = (\, e_{m+1} \quad e_{m+2} \; \ldots\ldots\ldots\ldots \; e_n \,)'$

$\beta_1 = (\, \phi_1 \ldots\ldots \phi_p \quad \Phi_1 \ldots\ldots \Phi_P \quad \alpha_{11} \ldots\ldots \alpha_{pP} \,)'$

$\beta 2 = (\, \omega_0 \ldots\ldots \omega_{h.} \quad \Omega_1 \ldots\ldots \Omega_H \quad \beta_{11} \ldots\ldots \beta_{hH} \,)'$

$\beta_3 = (\, \theta_1 \ldots\ldots \theta_q \quad \Theta_1 \ldots\ldots \Theta_Q \quad \gamma_{11} \ldots\ldots \gamma_{qQ} \,)'$

$$Y_{t-1} = \begin{bmatrix} y_m & \cdots\cdots\cdots\cdots & y_{n-1} \\ \cdots & \cdots\cdots\cdots\cdots & \cdots \\ y_{m-p+1} & \cdots\cdots\cdots\cdots & y_{n-p} \\ y_{m-s+1} & \cdots\cdots\cdots\cdots & y_{n-s} \\ \cdots & \cdots\cdots\cdots\cdots & \cdots \\ y_{m-Ps+1} & \cdots\cdots\cdots\cdots & y_{n-Ps} \\ y_{m-s} & \cdots\cdots\cdots\cdots & y_{n-s-1} \\ \cdots & \cdots\cdots\cdots\cdots & \cdots \\ y_{m-p-Ps+1} & \cdots\cdots\cdots\cdots & y_{n-p-Ps} \end{bmatrix}$$

## 2- THE SEASONAL ARMAX MODEL AND THE LIKELIHOOD FUNCTION

The seaonal ARMAX class of models my by defined as:

$$\phi(B)\,\Phi(B^s)\,y_t = \omega(B)\,\Omega(B^s)\,x_t + \theta(B)\,\Theta(B^s)e_t \qquad \text{.............................(2.1)}$$

where :

$y_t$ is the observation at time t , t = 1,2,3,........,n

$x_t$ 's are exogenous variables independent of $e_t$

$e_t$ 's are a sequence of independent random variables with mean zero and precision $\tau > 0$ , $\phi(B)$, $\omega(B)$ , $\theta(B)$ are polynomials in B of orders p,h and q respectivaly , $\Phi(B^s)$ , $\Omega(B^s)$ , and $\Theta(B^s)$ are polynomials in $B^s$ of orders P, H and Q respectivaly i.e

$$\phi(B) = 1-\phi_1 B - \phi_2 B^2 - \dots\dots - \phi_p B^p$$

$$\Phi(B^s) = 1-\Phi_1 B^s - \Phi_2 B^{2s} - \dots\dots - \Phi_P B^{Ps}$$

$$\omega(B) = \omega_0 - \omega_1 B - \omega_2 B^2 - \dots\dots - \omega_h B^h$$

$$\Omega(B^s) = 1-\Omega_1 B^s - \Omega_2 B^{2s} - \dots\dots - \Omega_H B^{Hs}$$

$$\theta(B) = 1-\theta_1 B - \theta_2 B^2 - \dots\dots - \theta_q B^q$$

$$\Theta(B^s) = 1-\Theta_1 B^s - \Theta_2 B^{2s} - \dots\dots - \Theta_Q B^{Qs}$$

and the backshift operator B is such that :

$$B^r y_t = y_{t-r}\ ,\ r=0,1,2,\dots \qquad \text{................(2.2)}$$

s is the number of the sesons.

The series generated by model (2.1) is assumed to be stationary - prehaps after an appropriate transformation - and there is no feedback from the output ( $y_t$) to the input ( $x_t$) variables i.e $x_t$'s are exogenous.

## 1- INTRODUCTION.

Forecasting the future values of an observed time series is very important in many areas , including business , economics , production planning , quality control , environmental studies and others. The time series analysis of the seasonal ARMAX model have been discussed in great detail - from a non-Bayesian viewpoint - in many text and surveys . Some of these text and surveys are : Gaynor and Krikpatrick (1994) , McGraw et al(1993), Lutkepohl(1993) , Mills(1991) , wei(1990) , Bierens(1987) , Spanos(1986) , Judge et al (1985) , Hendry et al(1984) , Deistler (1980), Box and Jenkins (1976) , Nerlove(1972), and others. On the other hand the Bayesian analysis of the ARMAX model is unknown because of the complexity of the likelihood function.

Through this paper a Bayesian procedure to forecast the future values of time series generated by seasonal autoregressive moving average model with exogenous variable , abbreviately seasonal ARMAX model , will be introduced. The predictive density of the future observations is the Bayesian tool to achieve this goal . The marginal predictive density of the first future observation will be derived . Also the conditional predictive densities of the observations at higher lags will be derived . In addition a highest predictive density regions (HPD regions ) for the future observations will be constructed .