

## Chương 9

# PHÂN TÍCH CHUỖI THỜI GIAN

## I. Chuỗi thời gian là gì ?

**Chuỗi thời gian** là số liệu của một biến số được quan sát và ghi nhận theo trình tự thời gian

**Phân tích chuỗi thời gian** là nghiên cứu hành vi, khuôn mẫu trong quá khứ của một biến số và sử dụng những thông tin này để dự đoán những giá trị hay những thay đổi của biến số đó trong tương lai

Có nhiều phương pháp để phân tích chuỗi thời gian nhưng thường gặp nhất là phương pháp Box - Jenkins

## II. Chuỗi thời gian dừng và không dừng

### 1. Định nghĩa chuỗi dừng và chuỗi không dừng

Chuỗi  $Y_t$  được gọi là **chuỗi dừng** (Stationary process) nếu kỳ vọng, phương sai và hiệp phương sai không đổi theo thời gian

$$E(Y_t) = \mu \text{ (const)}$$

$$\text{var}(Y_t) = \sigma^2 \text{ (const)}$$

$$\text{cov}(Y_t, Y_{t+k}) = \gamma_k \text{ (chỉ phụ thuộc vào } k, \text{ không phụ thuộc vào } t)$$

Chuỗi  $Y_t$  được gọi là **không dừng** (nonstationary process) nếu vi phạm ít nhất một trong 3 điều kiện trên

## II. Chuỗi thời gian dừng và không dừng

### 2. Kiểm định tính dừng của chuỗi

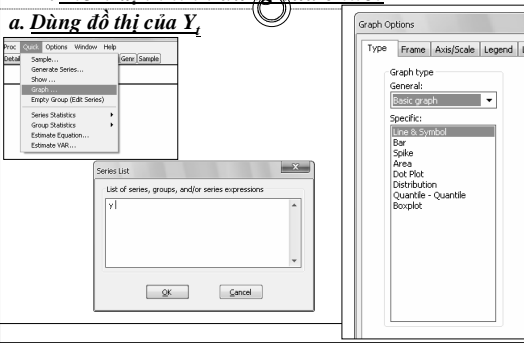
Khi khảo sát các **tính dừng** của chuỗi thời gian, người ta có thể dùng các cách sau :

- Dùng đồ thị của  $Y_t$  theo thời gian
- Dùng biểu đồ tự tương quan riêng của mẫu
- Dùng kiểm định bước ngẫu nhiên

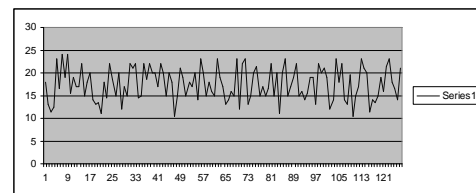
## II. Chuỗi thời gian dừng và không dừng

### 2. Kiểm định tính dừng của chuỗi

#### a. Dùng đồ thị của $Y_t$

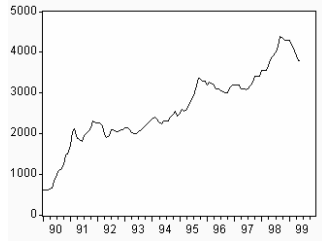


#### a. Dùng đồ thị của $Y_t$



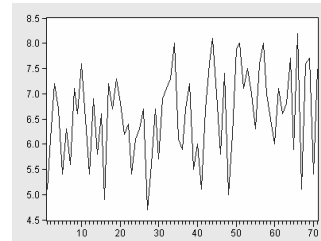
Đồ thị của một chuỗi dừng

a. Dùng đồ thị của  $Y_t$



Đồ thị của một chuỗi không dừng

a. Dùng đồ thị của  $Y_t$



Rất khó phân biệt

II. Chuỗi thời gian dừng và không dừng

2. Kiểm định tính dừng của chuỗi

b. Dùng đồ thị của tự tương quan mẫu

Khi khảo sát các tính dừng của chuỗi thời gian, người ta xây dựng hàm tự tương quan (Autoregressive Correlation Function - ACF) như sau :

$$ACF(k) = \rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\text{cov}(Y_t, Y_{t+k})}{\text{var}(Y_t)}$$

Ngoài ra còn tính hệ số tương quan riêng (partial autoregression correlation function - PACF)

II. Chuỗi thời gian dừng và không dừng

2. Kiểm định tính dừng của chuỗi

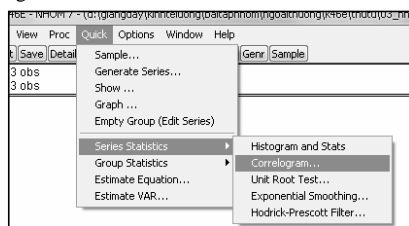
b. Dùng đồ thị của tự tương quan mẫu

Khi đó ACF đo hệ số tương quan giữa  $Y_t$  và  $Y_{t+k}$   
Trong khi PACF đánh giá mức độ tương quan giữa  $Y_t$  và  $Y_{t+k}$  khi bỏ qua các yếu tố trung gian

Các phần mềm thống kê ứng dụng và kinh tế lượng đều tính được ACF và PACF với các độ trễ khác nhau và vẽ lược đồ tương quan tương ứng

b. Dùng đồ thị của tự tương quan mẫu

Sau khi nhập dữ liệu, vào Quick/Series Statistics/Correlogram

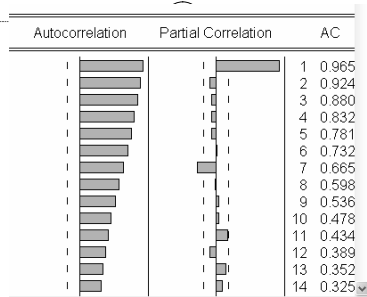


b. Dùng đồ thị của tự tương quan mẫu

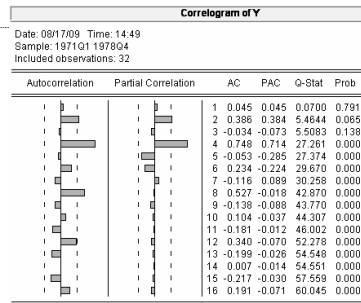
Date: 08/17/09 Time: 14:45  
Sample: 1 71  
Included observations: 71

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1.0000	1.0000	1.0000	1.0000	0.0000	0.0000
0.0200	0.0200	0.0200	0.0200	0.864	0.864
0.0050	0.0050	0.0050	0.0050	0.840	0.840
-0.0220	-0.0220	-0.0220	-0.0220	0.943	0.943
-0.1000	-0.1030	-0.1000	-0.1030	0.885	0.885
0.0087	0.0087	0.0087	0.0087	0.882	0.882
-0.0160	-0.0160	-0.0160	-0.0160	0.939	0.939
0.1950	0.1920	0.1950	0.1920	0.679	0.679
-0.0640	-0.0630	-0.0640	-0.0630	0.738	0.738
0.0560	0.0590	0.0560	0.0590	0.794	0.794
0.0010	-0.0000	0.0010	-0.0000	0.860	0.860
0.1030	0.1470	0.1030	0.1470	0.840	0.840
0.1370	0.0850	0.1370	0.0850	0.783	0.783
0.1320	0.1570	0.1320	0.1570	0.728	0.728
0.0800	0.0180	0.0800	0.0180	0.751	0.751
0.0140	0.0700	0.0140	0.0700	0.809	0.809
-0.0320	-0.0640	-0.0320	-0.0640	0.852	0.852

### b. Dùng đồ thị của tự tương quan mẫu



### b. Dùng đồ thị của tự tương quan mẫu



## II. Chuỗi thời gian dừng và không dừng

### 2. Kiểm định tính dừng của chuỗi

#### c. Dùng kiểm định bước ngẫu nhiên

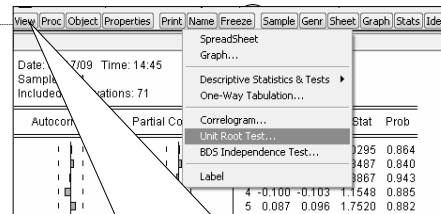
Kiểm định giả thiết sau với độ tin cậy  $(1 - \alpha)$

$H_0: Y_t$  là chuỗi không dừng

$H_1: Y_t$  là chuỗi dừng

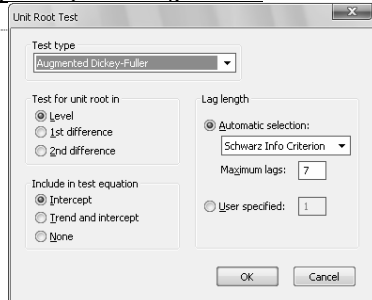
Ta dùng Eviews để thực hiện kiểm định này

#### c. Dùng kiểm định bước ngẫu nhiên



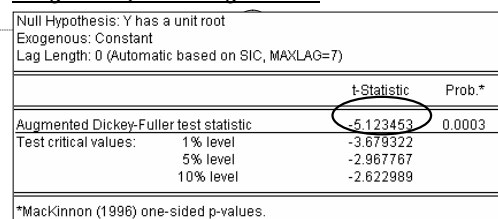
Trong lược đồ Correlogram, nhấn vào View \ Unit Root Test

#### c. Dùng kiểm định bước ngẫu nhiên



Chọn các thông số thích hợp

#### c. Dùng kiểm định bước ngẫu nhiên



Vì  $|\tau| > |\tau_\alpha|$  nên ta bác bỏ  $H_0$ , chấp nhận  $H_1$ , tức chuỗi đang xét là chuỗi dừng

## II. Chuỗi thời gian dừng và không dừng

### 2. Chuyển chuỗi không dừng thành chuỗi dừng

Giả sử ta có chuỗi thời gian  $Y_t$  là không dừng  
=> Lấy sai phân bậc nhất, bậc 2 hoặc bậc k thì sẽ được chuỗi dừng

Ký hiệu  $\Delta$  là toán tử sai phân

Sai phân cấp 1 :  $\Delta Y_t = V_t = Y_t - Y_{t-1}$

Sai phân cấp 2 :  $\Delta^2(Y_t) = \Delta(\Delta Y_t) = V_t - V_{t-1}$

.....

Sai phân cấp k :  $\Delta^k(Y_t)$

## II. Chuỗi thời gian dừng và không dừng

### 2. Chuyển chuỗi không dừng thành chuỗi dừng

$Y_t$  được gọi là liên kết bậc nhất, **I(1)**, nếu  $\Delta Y_t$  là chuỗi dừng  
 $Y_t$  được gọi là liên kết bậc hai, **I(2)**, nếu  $\Delta^2(Y_t)$  là chuỗi dừng

.....

$Y_t$  được gọi là liên kết bậc d, **I(d)**, nếu  $\Delta^d(Y_t)$  là chuỗi dừng

Nếu d = 0 thì  $Y_t$  đang xét là chuỗi dừng

Như vậy, **I(0)** tức là  $Y_t$  là chuỗi dừng

**I(1)** thì  $\Delta Y_t$  là chuỗi dừng ....

Để kiểm định sai phân bậc k của Y có là chuỗi dừng hay không thì ta kiểm định tương tự như phần trước

## III. Mô hình ARIMA , phương pháp Box-Jenkins

### 1. Mô hình AR, MA, ARMA và ARIMA

Nếu một chuỗi thời gian có tính dừng, ta có thể biểu diễn nó thành các mô hình sau :

#### a. Mô hình AR (Auto-Regressive)

Quá trình tự hồi quy bậc p có dạng :

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + U_t$$

Trong đó  $U_t$  là phần dư thỏa các giả thiết của mô hình cổ điển

## III. Mô hình ARIMA , phương pháp Box-Jenkins

### 1. Mô hình AR, MA, ARMA và ARIMA

#### b. Mô hình MA (Moving - Average)

$Y_t$  là quá trình trung bình trượt bậc p nếu  $Y_t$  có dạng :

$$Y_t = \theta_0 + U_t + \theta Y_{t-1} + \theta Y_{t-2} + \dots + \theta Y_{t-p}$$

## III. Mô hình ARIMA , phương pháp Box-Jenkins

### 1. Mô hình AR, MA, ARMA và ARIMA

#### c. Mô hình ARMA

Là kết hợp của mô hình AR và MA

Một quá trình ARMA(p,q) sẽ có p số hạng tự hồi quy (AR) và q số hạng trung bình trượt

Ví dụ nếu  $Y_t$  tuân theo quá trình ARMA(1,1)

$$Y_t = \theta_0 + \alpha_1 Y_{t-1} + \beta_0 U_t + \beta_1 U_{t-1}$$

## III. Mô hình ARIMA , phương pháp Box-Jenkins

### 1. Mô hình AR, MA, ARMA và ARIMA

#### c. Mô hình ARIMA

Giả sử  $Y_t$  là liên kết bậc d, **I(d)**, áp dụng mô hình ARMA(p,q) cho sai phân bậc d của  $Y_t$  thì ta có quá trình ARIMA(p,d,q)

Trong đó : p là bậc tự hồi quy

d bậc sai phân

q là bậc trung bình trượt

### III. Mô hình ARIMA , phương pháp Box-Jenkins

#### 1. Mô hình AR, MA, ARMA và ARIMA

##### c. Mô hình ARIMA

AR(p) có thể viết lại là ARIMA(p,0,0)

MA(q) có thể viết lại là ARIMA(0,0,q)

ARIMA(2,1,2) có phương trình hồi quy như sau

$$\Delta Y_t = \theta_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \beta_0 U_t + \beta_1 U_{t-1} + \beta_2 U_{t-2}$$

### III. Mô hình ARIMA , phương pháp Box-Jenkins

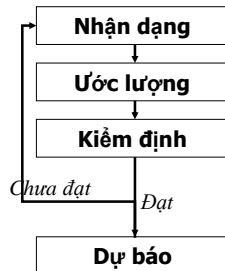
#### 2. Phương pháp Box-Jenkins

- Phương pháp BOX-JENKINS xem xét dữ liệu theo thời gian của một biến, xây dựng mô hình ARIMA thích hợp và dùng mô hình ARIMA đó để dự báo
- Phương pháp BOX-JENKINS được phát triển bởi 2 nhà thống kê G.E.P Box và G.M. Jenkins
- Theo Box-Jenkins thì mọi chuỗi thời gian đều có thể biểu diễn theo mô hình ARIMA(p,d,q) với các thông số phù hợp.

### III. Mô hình ARIMA , phương pháp Box-Jenkins

#### 2. Phương pháp Box-Jenkins

Phương pháp Box-Jenkins gồm 4 bước



#### 2. Phương pháp Box-Jenkins

##### a) Nhận dạng

Tìm p,d,q trong mô hình ARIMA(p,d,q)

p : dựa vào PACF

q : dựa vào ACF

d : số lần lấy sai phân để được chuỗi dừng

➢ Chọn mô hình AR(p) nếu PACF có giá trị cao tại độ trễ 1,2,...,p; giảm nhanh về 0 sau p và có ACF giảm dần về 0

➢ Chọn mô hình MA(q) nếu đồ thị ACF có giá trị cao tại độ trễ 1,2,...,q ; giảm nhanh về 0 sau q và có PACF giảm dần về 0

#### 2. Phương pháp Box-Jenkins

##### a) Nhận dạng

Loại mô hình	Dạng tiêu biểu của ACF	Dạng tiêu biểu của PACF
AR(p)	Suy giảm dần hay có dạng sóng hình sin giảm dần	Suy giảm đột ngột về 0 sau độ trễ p
MA(q)	Suy giảm đột ngột về 0 sau độ trễ q	Suy giảm dần hay có dạng sóng hình sin
ARMA(p,q)	Suy giảm nhanh về 0	Suy giảm nhanh về 0

#### 2. Phương pháp Box-Jenkins

##### a) Nhận dạng

Ví dụ, nếu Y<sub>t</sub> là dòng có đồ thị Correlogram như sau

Autocorrelation	Partial Correlation	AC
1	0.387	1
2	0.083	2
3	0.075	3
4	0.124	4
5	-0.071	5
6	-0.123	6
7	-0.043	7
8	0.081	8
9	-0.049	9
10	-0.042	10
11	0.068	11
12	0.112	12
13	-0.013	13
14	-0.129	14

AC giảm nhanh về 0 sau 1 độ trễ => q = 1

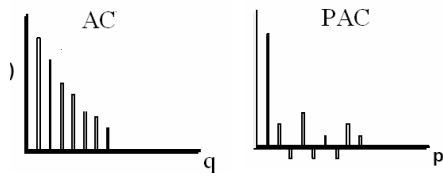
PAC giảm nhanh về 0 sau 1 độ trễ => p = 1

Mô hình thích hợp là ARIMA(1,0,1)

$$\text{dạng hàm: } Y_t = \theta_0 + \alpha_1 Y_{t-1} + \beta_0 U_t + \beta_1 U_{t-1}$$

## 2. Phương pháp Box-Jenkins

### a) Nhận dạng

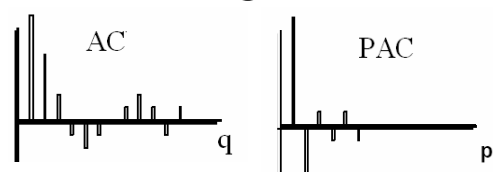


Mô hình thích hợp là gì ?

AC giảm dần, còn PAC giảm đột ngột về 0 sau 1 độ trễ

## 2. Phương pháp Box-Jenkins

### a) Nhận dạng

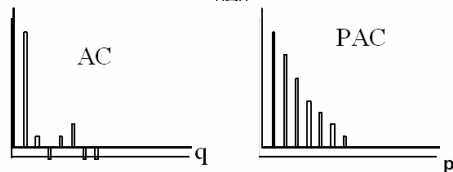


Mô hình thích hợp là gì ?

AC giảm dần và có dạng sóng hình sin, còn PAC giảm nhanh về 0 sau 2 độ trễ

## 2. Phương pháp Box-Jenkins

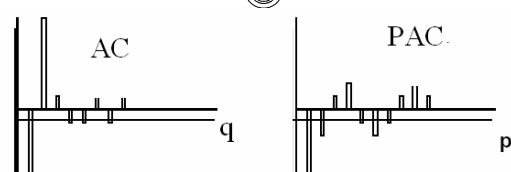
### a) Nhận dạng



Mô hình thích hợp là gì ?

## 2. Phương pháp Box-Jenkins

### a) Nhận dạng



Mô hình thích hợp là gì ?

Không có mô hình sai hay đúng, chỉ có mô hình chấp nhận được hay không mà thôi

## 2. Phương pháp Box-Jenkins

### b) Ước lượng

### Vấn dùng phương pháp OLS

Trong Eviews cũng vào Quick\ Estimate equation

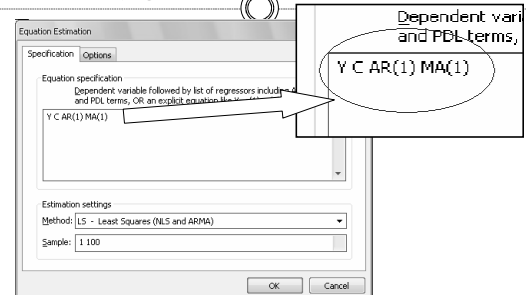
Ví dụ, trong phần nhận dạng ta xác định được  $Y_t$  tuân theo mô hình ARIMA(1,0,1) với phương trình :

$$Y_t = \theta_0 + \alpha_1 Y_{t-1} + \beta_0 U_t + \beta_1 U_{t-1}$$

Ta nhập các thông số như sau :

## 2. Phương pháp Box-Jenkins

### b) Ước lượng



## 2. Phương pháp Box-Jenkins

### b) Ước lượng

Kết quả

Dependent Variable: Y Method: Least Squares Date: 08/19/09 Time: 13:25 Sample (adjusted): 2 100 Included observations: 99 after adjustments Convergence achieved after 6 iterations MA Backcast: 1				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	133.0398	5.676676	23.43622	0.0000
AR(1)	-0.241553	0.136214	-1.773333	0.0793
MA(1)	0.724368	0.083988	7.707051	0.0000
R-squared	0.243235	Mean dependent var	133.3333	
Adjusted R-squared	0.227469	S.D. dependent var	46.31304	
S.E. of regression	40.70627	Akaike info criterion	10.28048	
Sum squared resid	159072.0	Schwarz criterion	10.35912	
Log likelihood	-505.8835	Hannan-Quinn criter.	10.31229	
F-statistic	15.42788	Durbin-Watson stat	1.794330	
Prob(F-statistic)	0.000002			
Inverted AR Roots	-.24			
Inverted MA Roots	-.72			

## 2. Phương pháp Box-Jenkins

### b) Ước lượng

Nếu là mô hình ARIMA(2,1,1), ta nhập thông số :

D(Y) C AR(1) AR(2) MA(1)

Kết quả

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.312863	3.519351	-0.088898	0.9294
AR(1)	-0.787319	0.189365	-4.157677	0.0001
AR(2)	-0.269574	0.093988	-2.868181	0.0051
MA(1)	0.845365	0.179223	3.600900	0.0005
R-squared	0.118444	Mean dependent var	0.412371	
Adjusted R-squared	0.090007	S.D. dependent var	45.48182	
S.E. of regression	43.38654	Akaike info criterion	10.41854	
Sum squared resid	175062.5	Schwarz criterion	10.52471	
Log likelihood	-501.2991	Hannan-Quinn criter.	10.46147	
F-statistic	4.165100	Durbin-Watson stat	2.083213	
Prob(F-statistic)	0.008152			

## 2. Phương pháp Box-Jenkins

### c) Kiểm định

- o Kiểm tra phần dư bằng thống kê Q (the Box-Pierce Q-Statistic)
- o Kiểm tra phần dư có phân phối chuẩn hay không ?

### o Kiểm tra phần dư bằng thống kê Q

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: Y  
Method: Least Squares  
Date: 08/19/09 Time: 22:56  
Sample (adjusted): 3 100  
Included observations: 98 after adjustments  
Convergence achieved after 1 iterations  
MA Backcast: 2

Representations  
Estimation Output  
Actual/Fitted/Residual  
ARMA Structure...  
Gradients and Derivatives  
Covariance Matrix

Coefficient Tests  
Residual Tests  
Stability Tests

Correlogram - Q-statistics  
Correlogram Squared Residuals  
Histogram - Normality Test  
Serial Correlation LM Test...  
Heteroskedasticity Tests...

Variable Coefficient Std. Error t-Statistic Prob.

C 13.0398 5.676676 23.43622 0.0000

AR(1) -0.241553 0.136214 -1.773333 0.0793

AR(2) 0.243958 0.119012 2.049690 0.0438

MA(1) 0.741919 0.148844 4.984900 0.0000

Lag Specification

Lags to include: 17

OK Cancel

### o Kiểm tra phần dư bằng thống kê Q

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	1	-0.001	-0.001	6.E-05	
2	0.013	0.013	0.0170		
3	0.045	0.045	0.2277		
4	0.002	0.002	0.2282	0.633	
5	-0.166	-0.168	3.1367	0.208	
6	-0.154	-0.161	5.6507	0.130	
7	-0.050	-0.052	5.9176	0.205	
8	-0.149	-0.138	8.3315	0.139	
9	-0.143	-0.147	10.594	0.102	
10	0.110	0.083	11.947	0.102	
11	0.153	0.134	14.593	0.068	
12	0.139	0.138	16.782	0.052	
13	0.032	-0.021	16.897	0.077	
14	0.150	0.057	19.537	0.052	
15	0.024	-0.004	19.603	0.075	
16	-0.033	-0.003	19.730	0.102	
17	0.001	0.038	19.731	0.139	

### o Kiểm tra phần dư có phân phối chuẩn hay không ?

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: Y  
Method: Least Squares  
Date: 08/19/09 Time: 22:56  
Sample (adjusted): 3 100  
Included observations: 98 after adjustments  
Convergence achieved after 1 iterations  
MA Backcast: 2

Representations  
Estimation Output  
Actual/Fitted/Residual  
ARMA Structure...  
Gradients and Derivatives  
Covariance Matrix

Coefficient Tests  
Residual Tests  
Stability Tests

Correlogram - Q-statistics  
Correlogram Squared Residuals  
Histogram - Normality Test  
Serial Correlation LM Test...  
Heteroskedasticity Tests...

Variable Coefficient Std. Error t-Statistic Prob.

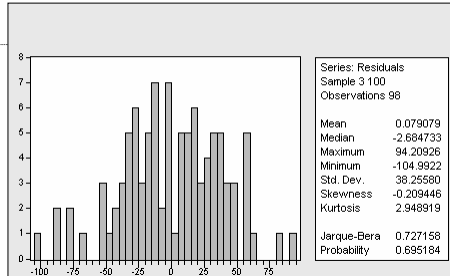
C 13.0398 5.676676 23.43622 0.0000

AR(1) -0.241553 0.136214 -1.773333 0.0793

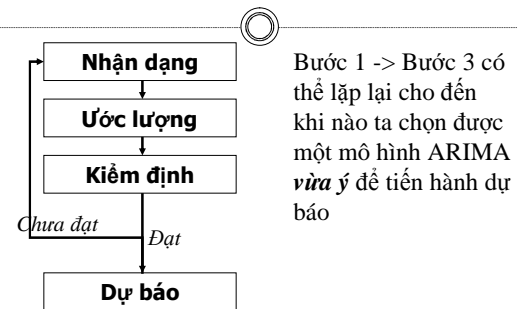
AR(2) 0.243958 0.119012 2.049690 0.0438

MA(1) 0.741919 0.148844 4.984900 0.0000

o **Kiểm tra phần dư có phân phối chuẩn hay không ?**



p-value >  $\alpha$  thì phần dư có phân phối chuẩn, mô hình chấp nhận được



Bước 1 -> Bước 3 có thể lặp lại cho đến khi nào ta chọn được một mô hình ARIMA *vừa ý* để tiến hành dự báo

## 2. Phương pháp Box-Jenkins

### d) Dự báo

Tiến hành dự báo với mô hình ARIMA vừa tìm được

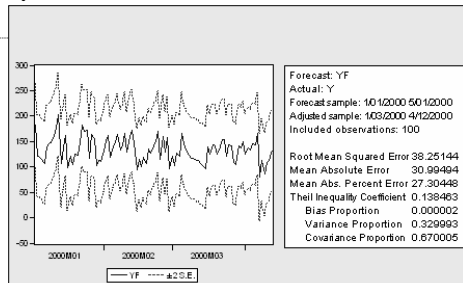
Từ cửa sổ phương trình hồi quy, nhấn nút **Forecast**

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

### d) Dự báo

Chọn thông số dự báo

### d) Dự báo



Đồ thị biểu diễn giá trị dự báo hiện ra  
Đồng đồ thị này lại

### d) Dự báo

#### Tiến hành dự báo khoảng

Từ màn hình workfile, nhấn nút **Genr**

Show Fetch Store Delete Genr Sample

Hoặc vào Quick\Generate Series

Quick Options Window Help  
Generate Series...



#### d) Dự báo

Tạo ra giá trị tin cậy dưới

Generate Series by Equation

Enter equation  
can\_duoi= yf-2\*yse

Sample  
1/01/2000 5/01/2000

OK Cancel

can\_duoi= yf-2\*yse

#### d) Dự báo

Tạo ra giá trị tin cậy trên

Generate Series by Equation

Enter equation  
can\_tren= yf+2\*yse

Sample  
1/01/2000 5/01/2000

OK Cancel

can\_tren= yf+2\*yse

#### d) Dự báo

Hiện thị giá trị dự báo

Proc Quick Options Window Help

Details Show Patch Store Delete Genr Sample

n -- 127 obs

Show

Objects to display in a single window  
y yf can\_duoi can\_tren

Enter one of the following  
- an Object or Object View  
- a Series Formula like LOG(X) or X+Y(-1)  
- a list of Series, Groups, and Formulas  
- a list of Graphs

OK

y yf can\_duoi can\_tren

#### d) Dự báo

obs	Y	YF	CAN_DUOI	CAN_TREN
4/01/2000	130.0000	124.7584	44.48461	205.0321
4/02/2000	140.0000	137.1248	57.11393	217.1357
4/03/2000	150.0000	132.6097	52.79010	212.4294
4/04/2000	160.0000	145.0744	65.15154	224.9973
4/05/2000	190.0000	143.3851	63.55985	223.2103
4/06/2000	60.00000	165.0810	84.59433	245.5677
4/07/2000	60.00000	80.40459	-5.157924	165.9671
4/08/2000	60.00000	113.1904	32.13714	194.2437
4/09/2000	60.00000	85.86071	3.759305	167.9621
4/10/2000	100.0000	108.6423	28.05581	189.2288
4/11/2000	120.0000	112.8255	32.13838	193.5125
4/12/2000	NA	132.5557	52.40293	212.7084
4/13/2000	NA	NA	NA	NA
4/14/2000	NA	NA	NA	NA
4/15/2000	NA	NA	NA	NA
4/16/2000	NA	NA	NA	NA
4/17/2000	NA	NA	NA	NA

Giá trị dự báo của kỳ tiếp theo