

TS. NGUYỄN DUY THỰC (CHỦ BIÊN)  
ThS. HOÀNG MẠNH HÙNG  
ThS. ĐÀM ĐÌNH MẠNH

*Dăng Chi Chuy*  
*Kế toán K36B*

**BÀI TẬP**  
**KINH TẾ LƯỢNG**

IN-PHOTO GIA RE  
NHƯ HÀ  
01.222.55.35.56- 01.222.55.35.58

**VỚI SỰ TRỢ GIÚP CỦA PHẦN MỀM EVIEWS 6**

*31 → 174*



NHÀ XUẤT BẢN KINH TẾ TP. HỒ CHÍ MINH

# MỤC LỤC

<b>LỜI NÓI ĐẦU .....</b>	<b>7</b>
<b>Chương 1: GIỚI THIỆU PHẦN MỀM EIEWS 6.0.....</b>	<b>9</b>
<b>1.1. Giới thiệu về phần mềm Eviews 6.0.....</b>	<b>9</b>
1.1.1. Cài đặt và chạy phần mềm Eviews 6.0 .....	10
1.1.2. Giao diện và các Menu chính.....	11
<b>1.2. Hướng dẫn nhập số liệu trong Eviews 6.0.....</b>	<b>13</b>
1.2.1. Tạo tập tin mới trực tiếp từ Eviews 6.0.....	13
1.2.2. Tạo tập tin mới bằng cách đọc từ phần mềm khác.....	17
<b>1.3. Trình bày dữ liệu trong Eviews 6.0.....</b>	<b>19</b>
1.3.1. Vẽ đồ thị.....	19
1.3.2. Thống kê mô tả.....	23
1.3.3. Đặt tên biến mới.....	25
1.3.4. Thay đổi cấu trúc dữ liệu .....	26
1.3.5. Lưu số liệu và mở file đã tạo .....	27
<b>1.4. Phép toán và hàm số trong Eviews 6.0 .....</b>	<b>27</b>
<b>1.5. Bài tập.....</b>	<b>28</b>
<b>Chương 2: MÔ HÌNH HỒI QUY .....</b>	<b>31</b>
<b>2.1. Mô hình hồi quy.....</b>	<b>31</b>
2.1.1. Mô hình hồi quy tuyến tính thông thường .....	31
2.1.2. Các dạng hồi quy thường gặp trong kinh tế.....	32
2.1.2.1. <i>Mô hình tuyến tính log</i> .....	32
2.1.2.2. <i>Mô hình semi – log</i> .....	32

2.1.2.3. Mô hình hồi quy đa thức .....	33
2.1.2.4. Mô hình nghịch đảo .....	34
2.1.2.5. Mô hình đi qua gốc toạ độ .....	34
<b>2.2. Ước lượng OLS.....</b>	<b>35</b>
<b>2.3. Hướng dẫn Eviews.....</b>	<b>36</b>
2.3.1. Ví dụ.....	36
2.3.2. Đọc bảng kết quả Eviews.....	39
2.3.3. Lưu kết quả Eviews.....	42
2.3.4. Xem giá trị thật, giá trị ước lượng và giá trị phần dư.....	42
<b>2.4. Bài tập.....</b>	<b>44</b>
<b>Chương 3: SUY DIỄN THỐNG KÊ.....</b>	<b>50</b>
<b>3.1. Các khoảng tin cậy và kiểm định giả thiết .....</b>	<b>50</b>
3.1.1. Khoảng tin cậy .....	50
3.1.2. Các kiểm định cơ bản cho mô hình hồi quy .....	52
3.1.2.1. Kiểm định về giá trị của hệ số hồi quy.....	52
3.1.2.2. Kiểm định về sự ràng buộc giữa các hệ số hồi quy..	53
3.1.2.3. Kiểm định thu hẹp hàm hồi quy .....	54
3.1.2.4. Kiểm định về sự phù hợp của hàm hồi quy.....	54
3.1.2.5. Kiểm định về phương sai nhiều.....	55
3.1.3. Dự báo giá trị của biến phụ thuộc .....	55
<b>3.2. Hướng dẫn Eviews.....</b>	<b>56</b>
3.2.1. Ví dụ 1 .....	56
3.2.2. Ví dụ 2.....	58
<b>3.3. Bài tập.....</b>	<b>69</b>

<b>Chương 4: PHÂN TÍCH HỒI QUY VỚI BIẾN ĐỊNH TÍNH.....</b>	<b>81</b>
<b>4.1. Mô hình hồi quy với biến giả .....</b>	<b>81</b>
<b>4.2. Mô hình với biến độc lập là định lượng và định tính....</b>	<b>82</b>
<b>4.3. Một số ứng dụng của biến giả .....</b>	<b>83</b>
4.3.1. Kiểm định sự thay đổi cấu trúc của mô hình.....	83
4.3.2. Biến giả trong phân tích mùa .....	84
<b>4.4. Hướng dẫn Eviews .....</b>	<b>85</b>
4.4.1. Kiểm định Chow.....	85
4.4.2. Thực hiện kiểm định thay đổi về cấu trúc với phương pháp biến giả .....	89
4.4.3. Ví dụ khác.....	91
<b>4.5. Bài tập .....</b>	<b>93</b>
<b>Chương 5: KIỂM ĐỊNH VÀ LỰA CHỌN MÔ HÌNH .....</b>	<b>97</b>
<b>5.1. Một số khuyết tật của mô hình hồi quy .....</b>	<b>97</b>
5.1.1. Phương sai sai số (PSSS) thay đổi .....	97
5.1.1.1. Kiểm định Glejer .....	97
5.1.1.2. Kiểm định Breusch – Pagan – Goldfrey (BPG, 1979).....	98
5.1.1.3. Kiểm định White .....	98
5.1.2. Tính chuẩn của sai số ngẫu nhiên (SSNN).....	100
5.1.3. Đa cộng tuyến.....	100
<b>5.2. Hướng dẫn Eviews.....</b>	<b>101</b>
5.2.1. Kiểm định Glejer.....	102
5.2.2. Kiểm định White và BPG.....	104

5.2.3. Ước lượng sai số chuẩn .....	106
5.2.4. Kiểm định tính chuẩn của sai số ngẫu nhiên .....	108
5.2.5. Kiểm định Ramsey – Reset .....	109
<b>5.3. Bài tập .....</b>	<b>111</b>
<b>Chương 6: MÔ HÌNH CHUỖI THỜI GIAN .....</b>	<b>122</b>
<b>6.1. Mô hình với số liệu chuỗi thời gian .....</b>	<b>122</b>
6.1.1. Mô hình .....	122
6.1.2. Hiện tượng tự tương quan .....	123
6.1.2.1. Kiểm định Durbin – Watson .....	123
6.1.2.2. Kiểm định Breusch – Godfrey .....	125
<b>6.2. Hướng dẫn Eviews .....</b>	<b>126</b>
<b>6.3. Bài tập .....</b>	<b>134</b>
<b>Chương 7: MỘT SỐ BÀI TẬP TỔNG HỢP .....</b>	<b>140</b>
<b>Phụ lục: CÁC BẢNG THỐNG KÊ .....</b>	<b>156</b>

## LỜI NÓI ĐẦU

Trong những năm gần đây, việc giảng dạy kinh tế lượng cho các chuyên ngành khác nhau tại các trường đại học được chú trọng và trở nên phổ biến. Sở dĩ như vậy vì không thể phủ nhận được khía cạnh ứng dụng quan trọng của kinh tế lượng trong thực tiễn, đặc biệt là khía cạnh phân tích hồi quy định lượng. Ngày nay, cùng với sự phát triển của máy vi tính, kinh tế lượng đã được áp dụng rộng rãi trong kinh tế, quản trị kinh doanh cũng như trong nhiều lĩnh vực khác.

Phương pháp giáo dục đại học đòi hỏi có sự tương tác giữa giảng viên và sinh viên trong quá trình giảng dạy, trong đó tập trung phát huy tính tích cực, chủ động của sinh viên, đòi hỏi mỗi sinh viên cần phải có kỹ năng tự học và tự nghiên cứu. Để đáp ứng được yêu cầu này thì việc cung cấp đầy đủ tài liệu học tập cho sinh viên là hết sức cần thiết. Đối với học phần kinh tế lượng, giáo trình và các tài liệu có liên quan còn rất thiếu hoặc chưa đủ, nhất là tài liệu về việc hướng dẫn các bài tập thông qua việc sử dụng các phần mềm chuyên dụng để phân tích. Bên cạnh đó thì thời lượng cho môn học cũng hạn chế nên nội dung thực hành ít được quan tâm. Hiểu biết về lý thuyết kinh tế lượng là cần thiết nhưng nếu không có kỹ năng sử dụng phần mềm trong thực hành thì không thể áp dụng được các phương pháp của kinh tế lượng để phân tích và dự báo kinh tế. Đứng trước vấn đề này, nhóm tác giả tiến hành biên soạn cuốn “Bài tập kinh tế lượng với sự trợ giúp của phần mềm Eviews 6” với hy vọng phần nào hỗ trợ giải quyết được khó khăn đó.

Cuốn sách bài tập gồm có 7 chương, chương 1 giới thiệu về phần mềm Eviews 6 và hướng dẫn những thao tác cơ bản ban đầu. Từ chương 2 đến chương 6 là hệ thống các bài tập phân tích hồi quy và vấn đề dự báo, trong đó tập trung chủ yếu vào việc sử

dụng phần mềm Eviews 6 để phân tích. Trong mỗi phần đều có tóm tắt những nội dung lý thuyết cơ bản và các ví dụ có hướng dẫn cụ thể để người học dễ dàng tự học tập và nghiên cứu. Chương 7 là hệ thống các bài tập tổng hợp bao gồm nhiều dạng bài tập khác nhau bao quát tất cả các nội dung trong chương trình môn học nhằm giúp cho người học củng cố và nắm vững hơn các vấn đề lý thuyết đã được học. Kèm theo cuốn sách là tất cả các file số liệu cho các ví dụ và bài tập được nêu trong cuốn sách bài tập này, đọc giả có thể download tại <http://www.fea.edu.vn>.

Hy vọng cuốn bài tập này sẽ là tài liệu bổ ích giúp cho các sinh viên, học viên cao học và cả những giảng viên quan tâm tiến nhanh hơn đến việc có thể sử dụng kinh tế lượng làm công cụ phục vụ cho việc nghiên cứu cũng như việc áp dụng cho thực tiễn cuộc sống.

Cuốn sách bài tập này được biên soạn lần đầu nên khó tránh khỏi những thiếu sót. Các tác giả rất mong nhận được sự góp ý và phê bình của bạn đọc để cuốn sách ngày càng có chất lượng tốt hơn.

*Tháng 01 năm 2015*  
Các tác giả

## **Chương 1:**

# **GIỚI THIỆU PHẦN MỀM EIEWS 6.0**

### **1.1. Giới thiệu về phần mềm Eviews 6.0**

EViews (kinh tế lượng xem) là một phần mềm thống kê cho Windows, được sử dụng chủ yếu cho chuỗi thời gian theo định hướng phân tích kinh tế. Nó được phát triển bởi công ty Quantitative Micro Software (QMS) bây giờ là một phần của HIS Inc. Phiên bản 1.0 được phát hành vào tháng Ba năm 1994, và thay thế MicroTSP. Phiên bản hiện tại của EViews là 8.0, được phát hành tháng ba năm 2013.

Eviews cung cấp các công cụ phân tích dữ liệu phức tạp, hồi quy và dự báo chạy trên Windows. Với Eviews ta có thể nhanh chóng xây dựng một mối quan hệ kinh tế lượng từ dữ liệu có sẵn và sử dụng mối quan hệ này để dự báo các giá trị tương lai. Eviews có thể hữu ích trong tất cả các loại nghiên cứu như đánh giá và phân tích dữ liệu khoa học, phân tích tài chính, mô phỏng và dự báo vĩ mô, dự báo doanh số, và phân tích chi phí. Đặc biệt, Eviews là một phần mềm rất mạnh cho phân tích dữ liệu thời gian cũng như chéo với cỡ mẫu lớn.

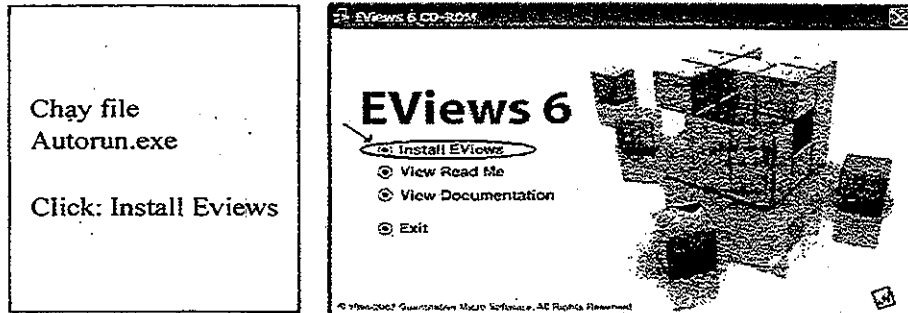
Eviews đưa ra nhiều cách nhập dữ liệu rất thông dụng và dễ sử dụng như nhập từ bàn phím, từ các file sẵn có dưới dạng excel hay text, dễ dàng mở rộng file dữ liệu có sẵn. Eviews trình bày các biểu mẫu, đồ thị, kết quả ấn tượng và có thể in trực tiếp hoặc chuyển qua các loại định dạng văn bản khác. Eviews giúp người sử dụng dễ dàng ước lượng và kiểm định các mô hình kinh tế lượng. Ngoài ra, Eviews còn giúp người nghiên cứu có thể xây dựng các file chương trình cho dự án nghiên cứu của mình. Eviews tận dụng các đặc điểm hiển thị của phần mềm Windows hiện đại nên rất thuận tiện cho người sử dụng như dùng chuột,

các thanh kéo, thay đổi giao diện, thoát, ...

### 1.1.1. Cài đặt và chạy phần mềm Eviews 6.0

#### Cài đặt phần mềm Eviews

Sau khi giải nén hết các file vào một thư mục nào đấy chúng ta tiến hành cài đặt theo các bước dưới đây:



Tiếp tục chọn Next → Yes → Next

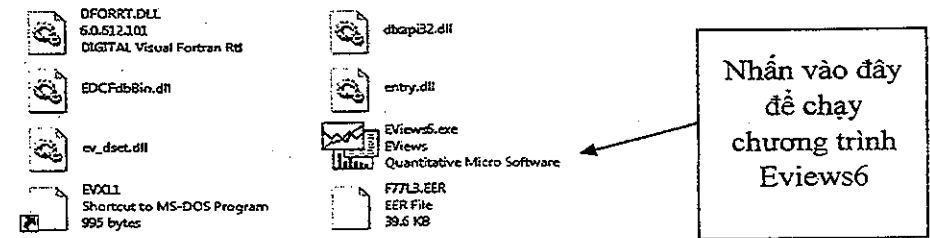
Tới bước này họ đòi serial thì các bạn điền vào là DEMO rồi tiếp tục nhấn OK và đợi cho cho trình cài đặt hoàn tất với bước cuối bạn nhấn Finish

Sau đó bạn hãy chạy file EViews6Patch\_060308.exe có trong thư mục Crack với việc cứ chọn Next liên tục cho tới cuối cùng nhấn Finish.

#### Khởi động phần mềm Eviews 6.0

Tệp chạy chương trình Eviews 6 là có biểu tượng là .

Nhấn vào biểu tượng của Eviews, cửa sổ chính của chương trình xuất hiện.



Trong Eviews, có thể xử lý đồng thời nhiều bộ số liệu, mỗi bộ số liệu trong một cửa sổ Tệp làm việc **Workfile**. Từ mỗi **Workfile**, thực hiện các thao tác với các đối tượng thông qua các cửa sổ riêng, một số đối tượng (*object*) chính:

- Series** xử lý cho từng biến số
- Group** xử lý một nhóm biến số
- Graph** vẽ các đồ thị
- Equation** hồi quy một phương trình
- System** hồi quy hệ phương trình

Một tập Cơ sở dữ liệu (*Database*) có thể gồm tất cả các đối tượng trên.

Trong tài liệu này, khi muốn xác định một nút trong cửa sổ nào, sẽ để cửa sổ trong cặp dấu ngoặc vuông [ ], ví dụ **[Workfile] Procs** là nút Procs của cửa sổ **Workfile**, **[Eviews]** là cửa sổ chính của Eviews. (→ : Thứ tự của thao tác.)

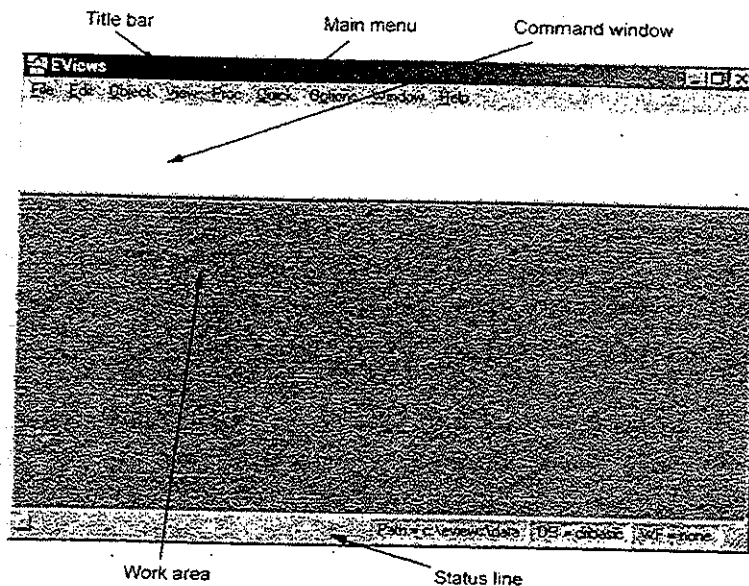
#### 1.1.2. Giao diện và các Menu chính

Cửa sổ chính của chương trình Eviews gồm các phần:

- ✓ Thanh menu: thực hiện các thao tác đã được chương trình định sẵn, tương tự như các chương trình chạy trong Windows.
- ✓ Cửa sổ lệnh: bên dưới thanh chức năng là nơi để viết lệnh trực tiếp. Có thể dùng chuột để kéo rộng cửa sổ lệnh tùy ý.

- ✓ Thanh chỉ dẫn: xác định đường dẫn đến tệp đang sử dụng.
- ✓ Các nút thu nhỏ, mở rộng cửa sổ, và thoát khỏi chương trình.

Khi khởi động cửa sổ Eviews 6 có dạng như sau:



Nguồn: *Eviews 6 Users Guide*, pp.16

Có thể không cần sử dụng chuột mà dùng bàn phím để chọn lựa các nút. Ấn và giữ phím Alt trên bàn phím, trên dòng task bar các lựa chọn sẽ tự động gạch chân các chữ cái. Khi đó, giữ phím Alt và nhấn phím tương ứng với chữ cái tương ứng sẽ cho kết quả giống như khi dùng chuột chọn nút đó.

**Ví dụ:** Khi giữ phím Alt, gõ phím F tương đương với nhấn chuột vào nút File; chữ E tương đương với nút Edit.

**Lưu ý:** - Ngăn cách phần nguyên và phần thập phân của một số, Eviews dùng dấu chấm “.”

- Phần mềm Eviews không phân biệt viết hoa và viết thường.
- Mức ý nghĩa thông thường là  $\alpha = 5\%$  cho các kiểm định, trong một số trường hợp lấy  $\alpha = 10\%$ .

## 1.2. Hướng dẫn nhập số liệu trong Eviews 6.0

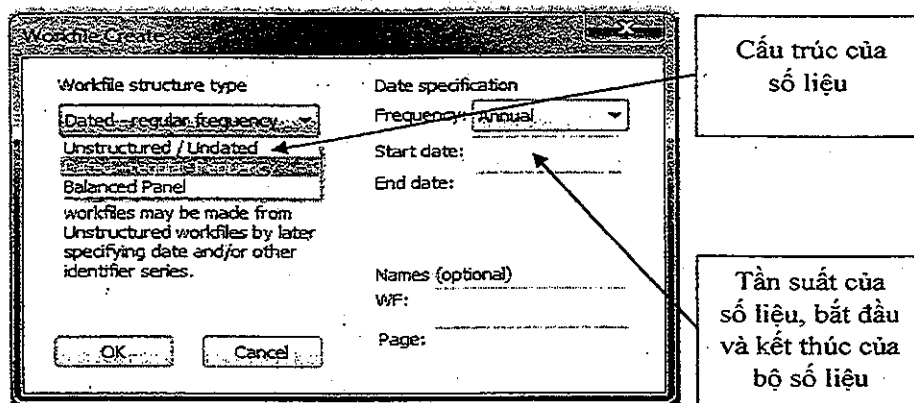
Tập tin làm việc của Eviews được gọi chung là Workfile (sau đây sẽ gọi là tập tin Eviews). Ở một cấp độ cơ bản, một tập tin Eviews đơn giản là một tập tin chứa các đối tượng của Eviews. Mỗi đối tượng bao gồm một tập hợp các thông tin có liên quan đến một lĩnh vực phân tích cụ thể ví dụ một chuỗi, một phương trình, hay một đồ thị. Làm việc trên Eviews chủ yếu liên quan đến các đối tượng chứa trong một tập tin Eviews. Cho nên trước hết cần tạo một tập tin mới hoặc mở một tập tin có sẵn. Mỗi tập tin Eviews chứa một hoặc nhiều trang. Mỗi trang chứa các đối tượng riêng. Trang được xem như một thư mục con hay tập tin phụ trong một tập tin. Lưu ý, một tập hợp các quan sát của một hoặc một số biến được gọi là bộ số liệu, trong đó mỗi quan sát có nhận dạng riêng.

Có nhiều cách tạo một tập tin mới. Việc đầu tiên trong tạo tập tin là xác định cấu trúc của tập tin. Có hai cách khác nhau. Thứ nhất là mô tả cấu trúc của tập tin Eviews. Theo cách này, Eviews sẽ tạo ra một tập tin mới để ta nhập dữ liệu vào một cách thủ công từ bàn phím hoặc copy và dán. Thứ hai là mở và đọc dữ liệu từ một nguồn bên ngoài (không thuộc định dạng Eviews) như Text, Excel, Stata. Thuận tiện hơn nhiều so với các phiên bản trước, Eviews 5 tự động phân tích nguồn dữ liệu, tạo một tập tin, và nhập dữ liệu.

### 1.2.1. Tạo tập tin mới trực tiếp từ Eviews 6.0

Để mô tả cấu trúc của tập tin Eviews, ta phải cung cấp cho Eviews các thông tin về số quan sát và các nhận dạng liên quan. Để tạo một tập tin mới ta thực hiện như sau:

[Eviews] → File → New → Workfile để vào bảng Workfile Create



Ở góc trái của hộp thoại là một hộp nhỏ để mô tả cấu trúc cơ bản của bộ dữ liệu. Ta có các lựa chọn sau đây:

**Unstructured/Undated:** nếu dữ liệu không có cấu trúc (dữ liệu chéo)

**Dated-Regular Frequency:** nếu dữ liệu có dạng ngày, tháng, năm

**Balanced Panel:** nếu dữ liệu ban đầu có dạng bảng

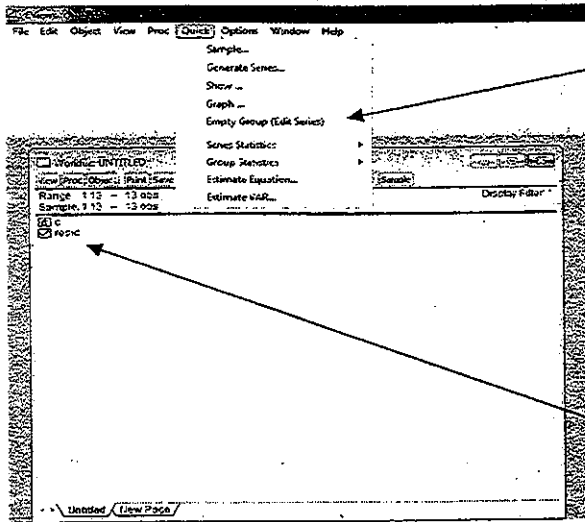
Sau khi ta đã xác định loại cấu trúc dữ liệu, Eviews sẽ tự động nhắc ta mô tả đặc điểm của bộ dữ liệu đó như tần suất, ngày bắt đầu, ngày kết thúc đối với loại dữ liệu thời gian; số quan sát đối với loại dữ liệu chéo; và tần suất, ngày bắt đầu, ngày kết thúc, và số quan sát tại mỗi thời điểm đối với loại dữ liệu bảng.

Quy tắc nhập ngày bắt đầu và kết thúc số liệu như sau:

Tần số	Start – End	Ý nghĩa
Annual (Năm) <i>yyyy</i>	Start: 1990 End: 2013	24 quan sát theo năm, từ năm 1990 đến năm 2013
Semi-annual (Nửa năm) <i>yyyy:h</i>	Start: 1990:1 End: 2013:2	48 quan sát theo nửa năm, từ nửa đầu năm 1990 đến nửa sau năm 2013
Quarterly (Quý) <i>yyyy:q</i>	Start: 1990:1 End: 2013:4	96 quan sát theo quý, từ quý 1 năm 1990 đến quý 4 năm 2013
Monthly (Tháng) <i>yyyy:mm</i>	Start: 1990:01 End: 2013:12	288 quan sát theo tháng, từ tháng 1 năm 1990 đến tháng 12 năm 2013
Weekly (Tuần) <i>mm/dd/yyyy</i>	Start: 01/01/2013 End: 12/01/2013	49 quan sát theo tuần, từ tuần có ngày 1 tháng 1 năm 2013 đến tuần có ngày 1 tháng 12 năm 2013
Daily [5day] (Ngày: tuần 5 ngày) <i>mm/dd/yyyy</i>	Start: 11/12/2013 End: 12/11/2013	22 quan sát theo ngày, từ ngày 12 tháng 11 năm 2013 đến ngày 11 tháng 12 năm 2013, không có ngày thứ 7 và chủ nhật
Daily [7day] (Ngày: tuần 7 ngày) <i>mm/dd/yyyy</i>	Start: 11/12/2013 End: 12/11/2013	29 quan sát theo ngày, từ ngày 12 tháng 11 năm 2013 đến ngày 11 tháng 12 năm 2013, có hai ngày cuối tuần
Unstructured/ Undated	Observations: 30	30 quan sát không theo thời gian (dạng số liệu chéo)



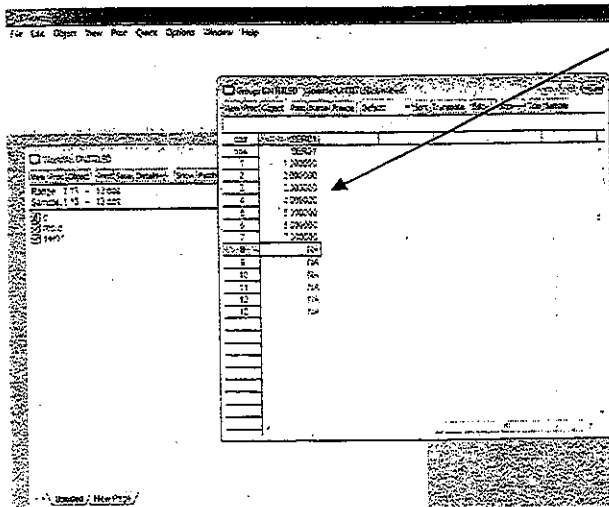
Sau đó chọn OK chúng ta sẽ tới một trang làm việc có dạng như sau:



Nhấn vào **Quick → Empty Group (Edit Series)** Để tạo một trang Group mới

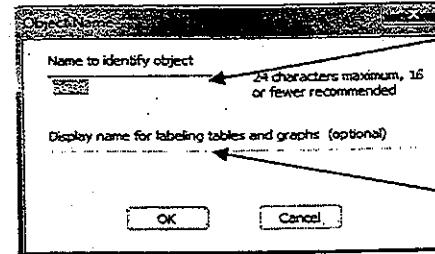
Vùng làm việc (luôn có sẵn 2 biến là C (hằng số) và resid (phần dư). Vì vậy khi đặt tên biến không được trùng với hai tên này

Lựa chọn Empty Group sẽ đưa ta đến hộp thoại sau:



Gõ trực tiếp số liệu vào vùng này hoặc copy từ Excel sang. Những phần dữ liệu thiếu sẽ có mặc định chữ NA. Eviews cũng mặc định đặt tên biến là SER01.. Ta có thể đổi tên trực tiếp trong trang Group.

Bạn có thể đổi tên biến số bằng cách nhấn chuột phải vào tên SER01 và chọn Rename



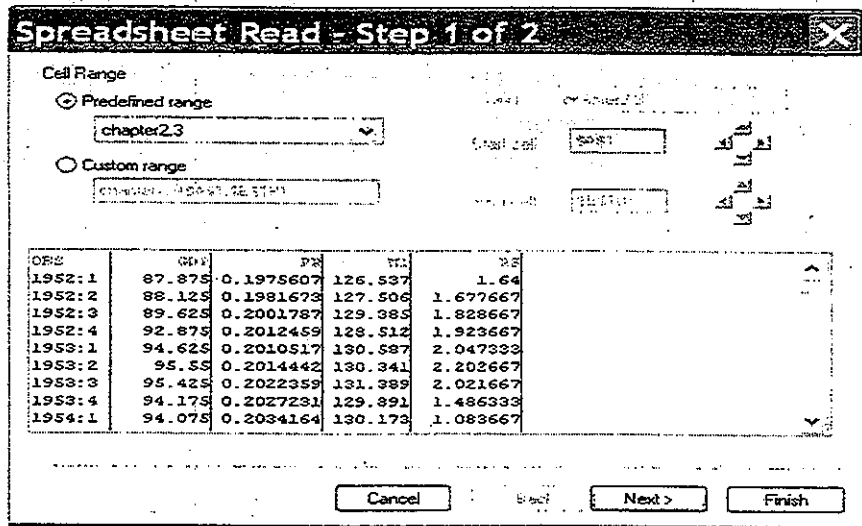
Đặt tên biến, không đặt tên tiếng Việt và ít hơn 16 ký tự là tốt nhất

Đặt nhãn cho biến số (nhãn sẽ hiển thị khi ta vẽ đồ thị hoặc trong bảng biểu thống kê)

### 1.2.2. Tạo tập tin mới bằng cách đọc từ phần mềm khác

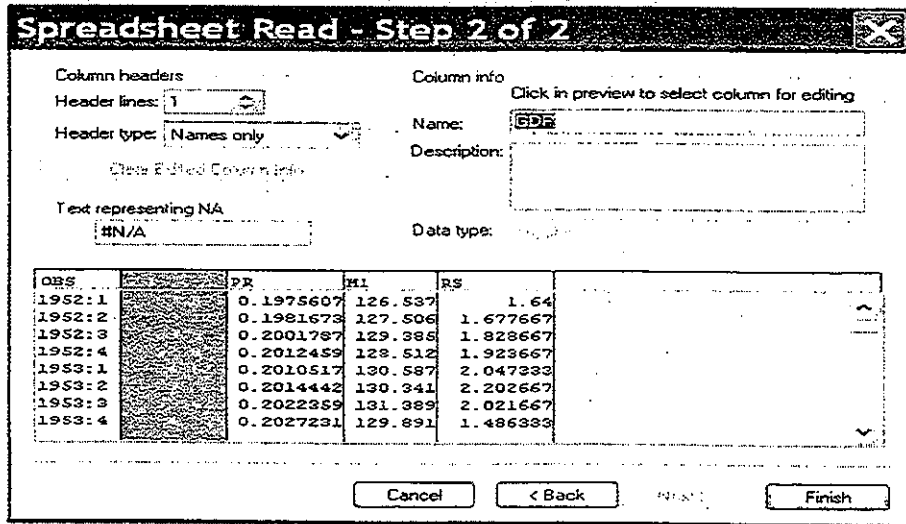
Ta có thể mở trực tiếp một nguồn dữ liệu bên ngoài như như cách mở một tập tin Eviews. Để mở một file bên ngoài, trước hết ta chọn **File/Open/Foreign Data as Workfile, ...** để đến hộp thoại **Open**, chọn **Files of type**, mở file cần chuyển sang tập tin Eviews, và thực hiện một số điều chỉnh nếu cần thiết. Xem ví dụ minh họa sau đây. Để mở và chuyển một tập tin nào đó sang Eviews, trước hết phải xác định thư mục thích hợp, rồi chọn tập tin (File name và Files of type) cần chuyển sang tập tin Eviews. Tuy nhiên, tập tin nguồn với định dạng khác nhau sẽ có một số khai báo riêng.

Thông thường nhất là sử dụng dữ liệu từ các tập tin Excel. Khi chọn và mở tập tin (ví dụ Chapter2.3.xls) Eviews sẽ thực hiện thông qua hai bước. Bước một, ta thấy xuất hiện hộp thoại **Spreadsheet Read** như sau:



Bước hai, giống như bước 3 ở tập tin dạng Text, Eviews sẽ đưa ra các lựa chọn để đọc dữ liệu và những thay đổi theo ý người sử dụng như đặt lại tên và nhãn của các biến.

Tuy nhiên, trong hầu hết các trường hợp người sử dụng chỉ cần chọn **Finish** để chấp nhận định dạng mặc định.



Vậy là chúng ta đã có dữ liệu để làm việc, việc đổi tên và gán nhãn hoàn toàn giống với cách thứ nhất.

### 1.3. Trình bày dữ liệu trong Eviews 6.0

#### 1.3.1. Vẽ đồ thị

Mô tả số liệu qua đồ thị và các thống kê đặc trưng cơ bản là xử lý ban đầu cần thiết đối với các biến số.

*Xem xét bộ số liệu ch2vd.wfl ta có các biến số:*

Y: tiêu dùng cuối cùng theo giá thực tế

X: GDP của Việt Nam tính theo giá thực tế

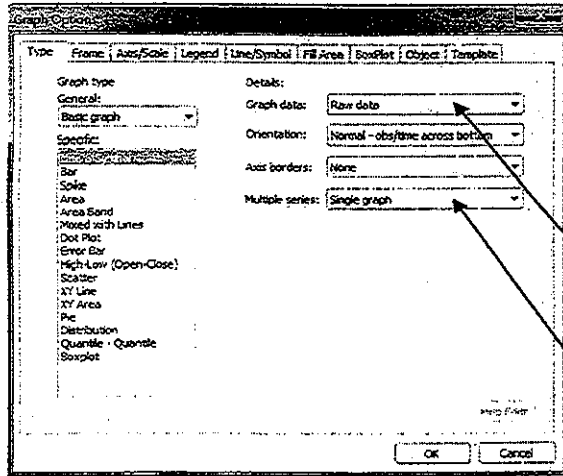
Số liệu lấy từ năm 1995 – 2011 (đơn vị: tỷ đồng).

Chọn các biến số chẳng hạn như X và Y, mở cửa sổ [Group] (Để chọn riêng các biến cách nhau, giữ phím Ctrl và chọn từng biến)

obs	Y	X
1995	167233.0	228892.0
1996	225231.0	272035.0
1997	250584.0	343523.0
1998	283444.0	361015.0
1999	301590.0	399942.0
2000	321853.0	441645.0
2001	342607.0	481295.0
2002	382137.0	535762.0
2003	445221.0	613443.0
2004	511221.0	715307.0
2005	584793.0	839211.0
2006	675916.0	974266.0
2007	809852.0	1143715
2008	1091875	1485038
2009	1206819	1658389
2010	1448901	1980914
2011	1794465	2535068

Các lệnh vẽ đồ thị nằm trong menu View ở đây hoặc Quick/Graph trên Menu chính

Chọn View → Graph... ta được hộp thoại sau:



Ta có thể chỉnh sửa các kiểu đồ thị, tên các trục, màu sắc ... trong hộp thoại này trước khi nhấn OK

Chọn dạng số liệu thể hiện trên đồ thị như: raw data, mean ...

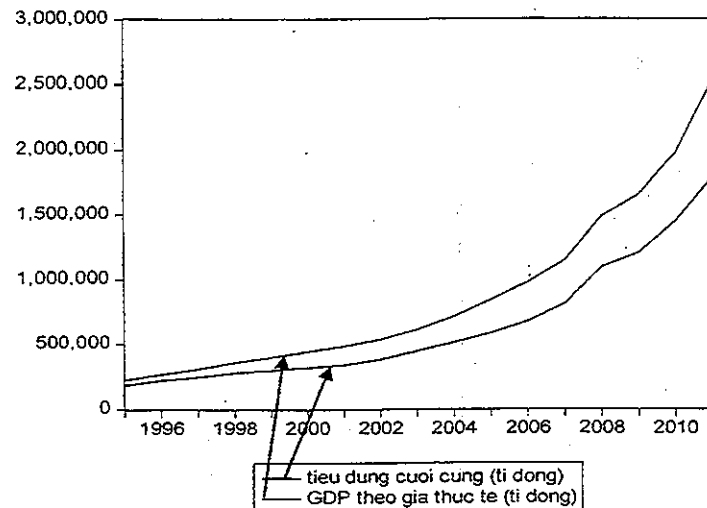
Chọn vẽ tất cả các biến trên một đồ thị hay hai đồ thị

- Vẽ đồ thị các biến theo quan sát

**Lưu ý:** dạng đồ thị đường (line) thường chỉ dùng cho số liệu chuỗi thời gian, số liệu chéo nên chọn đồ thị cột)

[Group] View → Graph → Line & Symbol

Kết quả cho đồ thị cột của X và Y theo quan sát trên cùng hệ tọa độ.

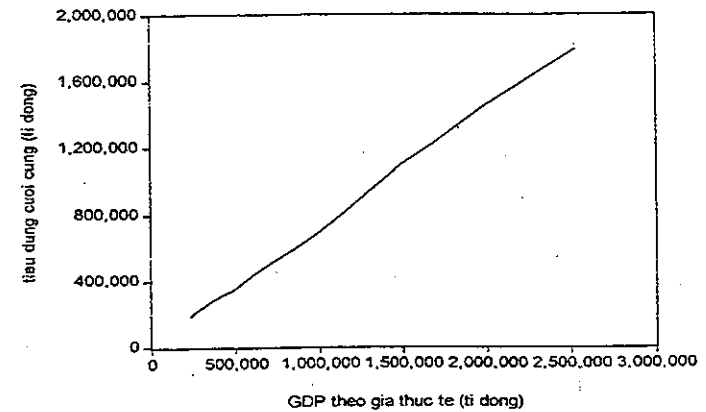


Hoặc mỗi biến một đồ thị thì trong mục **Multiple Series** chọn **Multiple Graphs**

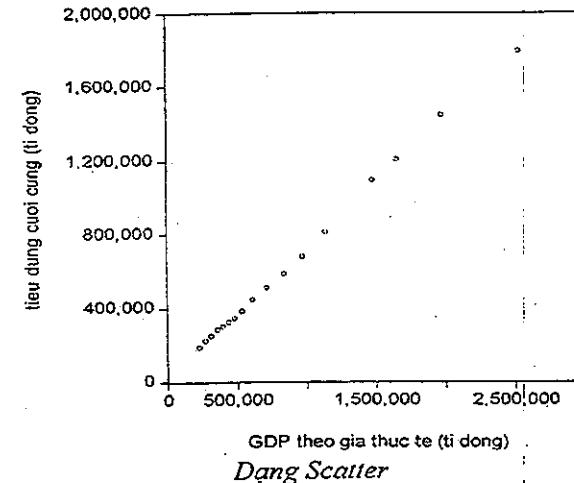
- Vẽ đồ thị của biến này theo biến kia, mỗi biến trên một trục tọa độ, lựa chọn:

[Group] View → Graph → XY line hoặc Scatter (phân tán)

Kết quả cho đồ thị điểm của biến Y trên trục tung và X trên trục hoành.



Dạng XY line

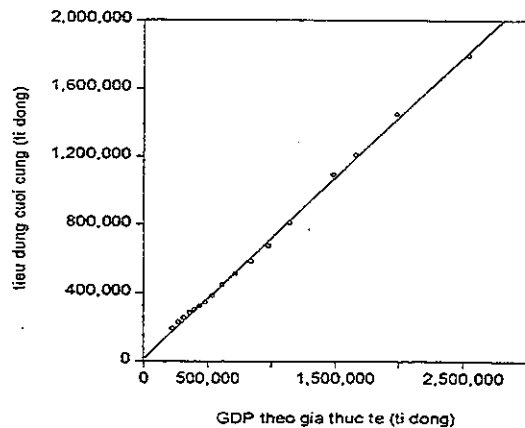


Dạng Scatter

**Lưu ý:** Eviews ngầm định biến xếp trước nằm trên trục hoành, biến xếp sau nằm ở trục tung.

- Xác định được hình ảnh của đường hồi quy bằng cách chọn:

[Group] View → Graph → Xyline thì trong mục Fit lines chọn Regression line



*Đồ thị có đường hồi quy tuyến tính*

#### ▪ Lưu đồ thị

Có thể lưu đồ thị để chèn vào các chương trình soạn thảo văn bản

Cửa sổ [Group] có đồ thị, nhấn tổ hợp Ctrl+C, mở cửa sổ [Graph Metafile], chọn OK, mở văn bản và nhấn Ctrl+V để dán đồ thị đã lưu.

#### ▪ Thay đổi định dạng đồ thị

Tại cửa sổ [Group] có đồ thị nhấp đúp chuột trái để vào lại cửa sổ [Graph Option] với các lựa chọn cho định dạng.

[Graph] Proc → Add text: Thêm dòng chữ vào đồ thị

[Graph] Proc → Save Graph: lưu đồ thị để chèn vào văn bản

[Graph] Name: lưu đồ thị dưới dạng một chủ thể, đặt tên.

Ngoài ra còn nhiều lựa chọn khác với đồ thị.

Bên cạnh cách chọn vẽ đồ thị từ cửa sổ [Graph], có thể có cách khác để vẽ đồ thị: Cửa sổ [Eviews] Quick → Graph: chọn loại đồ thị, và thứ tự của các biến để vẽ đồ thị.

#### Lưu ý:

Để chỉnh sửa và biên tập đồ thị, ta chọn Options hay nhấp đúp vào đồ thị. Eviews cho phép thay đổi hình nền, khung đồ thị, dạng đường đồ thị, đổi trục, đặt tên nhãn đồ thị, thay đổi font chữ,... Ngoài ra, Eviews cũng cho phép ta ghi chú dưới dạng text lên đồ thị, đặt tên đồ thị và lưu trong tập tin Eviews, hay có thể copy và dán dưới dạng văn bản. Ta cũng hay quan tâm đến các dạng biểu thị đồ thị bằng cách chọn Template, trong đó có nhiều sự lựa chọn rất thú vị. Nếu muốn lưu đồ thị (dạng một đối tượng trong tập tin Eviews), ta chọn Object/Freeze Output, rồi chọn Name để đặt tên đồ thị trong tập tin Eviews. Từ đồ thị, để trở lại bảng tính dữ liệu ta chọn View/SpreadSheet.

#### 1.3.2. Thông kê mô tả

- [Group] View → Descriptive Stats → Common Sample : Các thống kê mô tả từng biến
- [Group] View → Correlations → Common Sample: Hệ số tương quan các biến.
- [Group] View → Covariances → Common Sample: Phương sai - Hiệp phương sai
- (Có thể dùng [Eviews] Quick → Group Statistic → Descriptive Stats → Common Sample)

Với bộ số liệu **ch2vd.wf1** ở trên ta có thể có các thống kê mô tả sau:

	X	Y
Mean	881147.2	638932.5
Median	613443.0	445221.0
Maximum	2535008.	1794465.
Minimum	228892.0	187233.0
Std. Dev.	670123.0	475885.0
Skewness	1.158250	1.173037
Kurtosis	3.319154	3.240148
Jarque-Bera	3.873186	3.939564
Probability	0.144194	0.139487
Sum	14979503	10861853
Sum Sq. Dev.	7.19E+12	3.62E+12
Observations	17	17

Mean: Giá trị trung bình

Median: Trung vị

Maximum: Giá trị lớn nhất

Minimum: Giá trị nhỏ nhất

Std. Dev.: Độ lệch chuẩn

Skewness: Hệ số bất đối xứng

Kurtosis: Hệ số nhọn

Jarque-Bera: Hệ số Jarque-Bera

Probability: Xác suất của hệ số Jarque-Bera

Sum: Tổng các giá trị của biến

Sum Sq. Dev.: Tổng bình phương các độ lệch giữa các giá trị quan sát của từng biến với giá trị trung bình của nó

Observations: Số quan sát

### 1.3.3. Đặt tên biến mới

Một số biến mới có thể được đặt từ các biến có sẵn, hoặc một số biến đặc biệt có thể đặt trực tiếp. Việc đặt biến thực hiện trong cửa sổ lệnh. Một số trường hợp đặt biến ví dụ:

**Genr EX2 = EX^2** Biến *EX2* bằng bình phương của *EX*

**Genr DEX = D(EX)** Biến *DEX* bằng sai phân bậc nhất của *EX*

**Genr LEX = LOG(EX)** Biến *LEX* bằng logarit cơ số tự nhiên của *EX*

**Genr T = @TREND()** *T* là biến xu thế thời gian với quan sát đầu bằng 0

Có thể đặt biến mới bằng cách tạo biến trống, rồi gán giá trị biến bằng một công thức tương ứng trong cửa sổ [Series] Procs → Generate by equation). Khi muốn xóa bớt biến số, có nhiều cách: Chọn biến → chuột phải → Delete; hoặc gõ trong cửa sổ lệnh: Detele [tên biến])

Bên cạnh các biến đã nhập số liệu là X2, X3 và Y, có thể đặt các biến mới từ các biến đã có hoặc nhập biến số mới.

Ví dụ: cần đặt biến mới tổng thu nhập:  $Z = X2 + X3$

[Eviews] Quick → Generates Series: mở cửa sổ [Generate Series by Equation]

Tại ô Enter Equation gõ:  $Z = X2 + X3$

Tại ô Sample, ngầm định mẫu mà lệnh đặt biến có tác dụng là toàn bộ.

Khi cần có thể thay đổi mẫu này. Nhấn OK để chấp nhận.

### 1.3.4. Thay đổi cấu trúc dữ liệu

[Group] Procs → Sample	Xác định lại thời kỳ mẫu
[Group] Procs → Change Workfile Range	Thêm hoặc bớt quan sát
[Group] Procs → Generate Series	Đặt lại biến mới bởi một phương trình
[Group] Procs → Import	Nhập số liệu từ file có sẵn
[Group] Objects → New Objects → ...	Tạo thêm một đối tượng
[Group] Objects → Store to DB	Lưu vào Database (nếu đã tạo Database)
[Group] Objects → Name	Đặt tên cho Group $\square$ trong Workfile

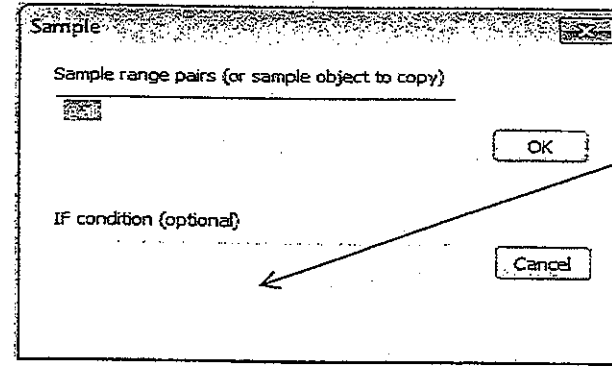
(Có thể dùng [Group] Name trực tiếp)

Trong nhiều trường hợp, do giới hạn hoặc qui định kích thước mẫu nghiên cứu mà ta cần phải xác định lại mẫu trước khi thực hiện.

Để định lại mẫu, từ cửa sổ Workfile chọn Sample ta được:

@all: sử dụng toàn bộ dãy dữ liệu của tập tin, nếu chỉ sử dụng một số quan sát từ m đến k thì ta thay @all bằng m k

IF condition (optional): điều kiện mà mẫu cần thỏa mãn



Nhập điều kiện mẫu cần thỏa mãn

### 1.3.5. Lưu số liệu và mở file đã tạo

Thực hiện xong một vài thao tác, muốn lưu lại bài làm ta thực hiện như sau:

[Eviews] File → Save As → Đặt tên cho bộ số liệu, phân mở rộng ngầm định là .wfl

Hoặc [Eviews] File → Save nếu muốn lưu đè vào file đã tạo.

**Lưu ý:** Tên file không quá 256 ký tự, không nên sử dụng tên tiếng Việt (dễ bị lỗi khi mở file)

Việc mở file số liệu đã có cũng hoàn toàn đơn giản:

[Eviews] File → Open → Eviews Workfile...

### 1.4. Phép toán và hàm số trong Eviews 6.0

Ngoài các phép toán cơ bản: cộng [+], trừ [-], nhân [\*], chia [/], lũy thừa [^], các phép toán trong Eviews như sau:

Tên	Ý nghĩa	Ví dụ	Kết quả
LOG	Logarit tự nhiên	LX = LOG(X)	$LX_i = \ln(X_i)$
EXP	Hàm mũ cơ số tự nhiên	EX = EXP(X)	$EX = e^x$
ABS	Giá trị tuyệt đối	AX = ABS(X)	$AX_i =  X_i $
SQR	Căn bậc hai	SX = SQR(X)	$SX_i = \sqrt{X_i}$
SIN	Hàm sin	SIN(X)	$Sin(X)$
COS	Hàm cosin	COS(X)	$Cosin(X)$
(- k)	Trễ bậc k	X4 = X(-4)	$X4_i = X_{i-4}$
D(X)	Sai phân bậc 1	DX = D(X)	$DX_i = X_i - X_{i-1}$
D(X,k)	Sai phân bậc k	DKX = D(X,k)	$D^k(X) = X_i - X_{i-k}$

Các phép so sánh trong Eviews được qui ước như sau: lớn hơn (>), nhỏ hơn (<), lớn hơn hoặc bằng (>=), nhỏ hơn hoặc bằng (<=) và bằng (=).

### 1.5. Bài tập

Cho bảng số liệu sau với:

Year: Năm quan sát từ 1959 đến 2001

M2: lượng cung tiền đã hiệu chỉnh yếu tố mùa vụ (đơn vị: tỷ USD)

GDP: thu nhập quốc dân đã hiệu chỉnh yếu tố mùa vụ (đơn vị: tỷ USD)

Interest: lãi suất chiết khấu (đơn vị : %)

Year	M2	GDP	interest	Year	M2	GDP	interest
1959	297.8	2319.0	3.36	1981	1754.9	5021.0	13.42
1960	312.4	2376.7	3.53	1982	1909.8	4919.3	11.01
1961	335.5	2432.0	3.00	1983	2125.9	5132.3	8.50
1962	362.7	2578.9	3.00	1984	2309.6	5505.2	8.80
1963	393.2	2690.4	3.23	1985	2494.9	5717.1	7.69
1964	424.7	2846.5	3.55	1986	2731.6	5912.4	6.32
1965	459.2	3028.5	4.04	1987	2830.6	6113.3	5.66
1966	480.2	3227.5	4.50	1988	2993.8	6368.4	6.20
1967	524.8	3308.3	4.19	1989	3157.4	6591.8	6.93
1968	566.8	3466.1	5.17	1990	3276.8	6707.9	6.98
1969	587.9	3571.4	5.87	1991	3376.1	6676.4	5.45
1970	626.4	3578.0	5.95	1992	3430.3	6880.0	3.25
1971	710.1	3697.7	4.88	1993	3483.0	7062.6	3.00
1972	802.1	3898.4	4.50	1994	3496.0	7347.7	3.60
1973	855.3	4123.4	6.45	1995	3638.7	7543.8	5.21
1974	901.9	4099.0	7.83	1996	3811.3	7813.2	5.02
1975	1016.0	4084.4	6.25	1997	4028.0	8159.5	5.00
1976	1151.7	4311.7	5.50	1998	4380.5	8508.9	4.92
1977	1269.9	4511.8	5.46	1999	4650.3	8856.5	4.62
1978	1365.6	4760.6	7.46	2000	4936.0	9224.0	5.73
1979	1473.3	4912.1	10.29	2001	5454.8	9333.8	3.40
1980	1599.4	4900.9	11.77				

Nguồn: GDP - Bureau of Economic Analysis, U.S. Department of Commerce (<http://www.bea.doc.gov/bea/dn1.htm>). M2, int - Federal reserve Board - <http://www.federalreserve.gov>

### Yêu cầu:

1. Nhập số liệu vào Eviews và lưu thành file **ch1.1.wfl**
2. Vẽ đồ thị của các biến số theo thời gian và nhận xét
3. Tính toán các thống kê mô tả
4. Cho biết hệ số tương quan giữa các biến M2, GDP, Interest
5. Tạo biến xu thế T với  $T = 1$  vào năm 1959
6. Tạo biến mới  $GDP^2$  và đặt tên hợp lý
7. Tạo các biến  $\ln GDP$ ,  $\ln M2$ ,  $\ln Interest$  và vẽ đồ thị. So sánh các dạng đồ thị này với đồ thị của biến gốc tại câu 2.

## Chương 2: MÔ HÌNH HỒI QUY

### 2.1. Mô hình hồi quy

#### 2.1.1. Mô hình hồi quy tuyến tính thông thường

Mô hình hồi quy tổng quát k biến là mô hình trong đó có một biến phụ thuộc và  $(k - 1)$  biến độc lập.

Cho biến phụ thuộc là:  $Y$

Các biến độc lập là:  $X_2, X_3, \dots, X_k$

Các dạng của mô hình hồi quy k biến như sau:

Trên toàn bộ tổng thể thống kê của hiện tượng nghiên cứu:

$$\begin{aligned} \bullet \text{ PRF: } E(Y / X_{2i}, \dots, X_{ki}) &= \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki} \\ E(Y_i) &= \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki} \end{aligned} \quad (2.1)$$

Dạng ngẫu nhiên của PRF

$$Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i \quad (2.2)$$

(Hoặc  $E(Y_i) = \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki}$ ;  $Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i$ , nếu là số liệu chuỗi thời gian).

\* Ý nghĩa các hệ số hồi quy:

+  $\beta_1$  được gọi là hệ số chặn, nó là giá trị trung bình của biến phụ thuộc  $Y$  khi các biến độc lập trong mô hình nhận giá trị bằng 0. Ý nghĩa của hệ số này thường ít được quan tâm.

+  $\beta_2, \dots, \beta_k$  được gọi là các hệ số góc hay các hệ số hồi



quy riêng. Hệ số hồi quy  $\beta_j, j = \overline{2, k}$  có ý nghĩa là khi các yếu tố  $X_s (s \neq j)$  không đổi, nếu  $X_j$  tăng (giảm) 1 đơn vị thì trung bình  $Y$  tăng (giảm)  $\beta_j$  đơn vị.

Với mẫu  $W = \{(X_{2i}, X_{3i}, \dots, X_{ki}, Y_i); i = \overline{1, n}\}$

$$\bullet \text{ SRF: } \hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \dots + \hat{\beta}_k X_{ki} \quad (2.3)$$

Với  $\hat{\beta}_j$  là ước lượng của  $\beta_j (j = \overline{1, k})$  từ mẫu trên.

Dạng ngẫu nhiên của SRF

$$Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + e_i \quad (2.4)$$

Các dạng mô hình trên gọi là mô hình hồi quy tuyến tính thông thường (tính tuyến tính chỉ giới hạn trên các tham số hồi quy  $\beta_j$ ).

### 2.1.2. Các dạng hồi quy thường gặp trong kinh tế

Với các biến số kinh tế và xã hội, mối quan hệ giữa chúng thể hiện bằng các dạng hàm hồi quy khác nhau tùy theo tính chất và đặc trưng riêng.

#### 2.1.2.1. Mô hình tuyến tính log

$$\ln Y_i = \beta_1 + \beta_2 \ln X_{2i} + \dots + \beta_k \ln X_{ki} + u_i$$

Nếu  $X_j$  tăng (giảm) 1% và các yếu tố khác trong mô hình không đổi thì trung bình  $Y$  tăng (giảm)  $\beta_j\%$ . Các hệ số góc này còn gọi là hệ số co giãn. Và dạng hàm tuyến tính log này thường được dùng để mô tả các mối quan hệ trong đó hệ số co giãn là không đổi.

#### 2.1.2.2. Mô hình semi-log

Mô hình log-lin:  $\ln Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i$

Mô hình lin-log:  $Y_i = \beta_1 + \beta_2 \ln X_{2i} + \dots + \beta_k \ln X_{ki} + u_i$

Các hệ số  $\beta_j, j = \overline{2, k}$  trong mô hình log-lin được giải thích là khi các yếu tố  $X_s (s \neq j)$  không đổi, nếu  $X_j$  tăng (giảm) 1 đơn vị thì trung bình  $Y$  tăng (giảm)  $100 \cdot \beta_j\%$ .

Các hệ số  $\beta_j, j = \overline{2, k}$  trong mô hình lin-log được giải thích là khi các yếu tố  $X_s (s \neq j)$  không đổi, nếu  $X_j$  tăng (giảm) 1% thì trung bình  $Y$  tăng (giảm)  $\beta_j/100$  đơn vị.

Các mô hình được trình bày trên đây là thuần log-log hoặc semi-log. Tuy nhiên trên thực tế ta có thể gặp các mô hình hỗn hợp, chẳng hạn mô hình sau:

$$Y_i = \beta_1 + \beta_2 \ln X_i + \beta_3 X_{3i} + u_i$$

*Vậy khi nào nên lựa chọn dạng log cho biến số?*

- Dạng hàm log thường được lựa chọn khi có gợi ý từ lý thuyết kinh tế về mối quan hệ giữa các biến số.
- Dạng hàm log cũng thường được dùng khi các biến số đều nhận giá trị dương như dân số, GDP, lao động, tiền lương hoặc các biến số mà phân phối có đuôi lệch như thu nhập, mức lương ... Việc lấy logarit cho các biến số này làm cho phân phối của nhiều gần với phân phối chuẩn hơn.

Dạng hàm này có ưu thế là các kết quả ước lượng không phụ thuộc vào đơn vị đo của biến số.

#### 2.1.2.3. Mô hình hồi quy đa thức

$$Y_i = \beta_1 + \beta_2 X_i + \beta_3 X_i^2 + \dots + \beta_k X_i^{k-1} + u_i$$

Ta thấy trong những hàm hồi quy đa thức này chỉ có một biến độc lập ở vế phải nhưng nó xuất hiện với những lũy thừa khác nhau khiến cho chúng trở thành mô hình hồi quy bội. Những mô hình này vẫn là mô hình tuyến tính theo tham số nên chúng có thể ước lượng bằng phương pháp OLS thông thường.

Mô hình hồi quy đa thức cũng thường được dùng để nghiên cứu hàm chi phí hoặc tiền lương.

#### 2.1.2.4. Mô hình nghịch đảo

Mô hình sau đây gọi là mô hình nghịch đảo:

$$Y_i = \beta_1 + \beta_2 \frac{1}{X_i} + u_i$$

Đặc điểm của mô hình này là khi biến X tăng vô hạn thì  $1/X$  tiến về 0, khi đó biến phụ thuộc Y tiến về  $\beta_1$ , giá trị này gọi là tiệm cận ngang. Với hệ số góc  $\beta_2$  nếu nó dương thì Y là hàm giảm theo X, ngược lại nếu nó âm thì Y là hàm tăng theo X và  $E(Y_i) = 0$  thì  $X = -\beta_2/\beta_1$ .

Một vài trường hợp sử dụng mô hình nghịch đảo:

- ✓ Xét mối quan hệ giữa chi phí sản xuất cố định trung bình (AFC) và sản lượng. Khi sản lượng tăng thì chi phí sản xuất cố định trung bình trên một sản phẩm có khuynh hướng giảm dần nhưng không vượt qua mức tối thiểu.
- ✓ Mối quan hệ giữa tỷ lệ thay đổi tiền lương (Y) và tỷ lệ thất nghiệp (X) trong kinh tế vĩ mô được biểu diễn bằng đường cong Phillips.
- ✓ Đường chi tiêu Engel biểu diễn mối quan hệ chi tiêu của người tiêu dùng cho một hàng hoá (Y) với tổng chi tiêu hay thu nhập (X) của người đó.

#### 2.1.2.5. Mô hình đi qua gốc toạ độ

$$Y_i = \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i$$

Đây là dạng mô hình không có hệ số chặn. Thông thường mô hình này rất hạn chế sử dụng trừ khi ta có cơ sở lý thuyết hoặc có trước kinh nghiệm tốt, ví dụ mô hình định giá tài sản vốn CAPM của lý thuyết danh mục đầu tư hiện đại.

## 2.2. Ước lượng OLS

Xét mô hình hồi quy:  $Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i$

Từ mẫu:  $W = \{(X_{2i}, X_{3i}, \dots, X_{ki}, Y_i); i \in \overline{1, n}\}$  ta có dạng hàm SRF:

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \dots + \hat{\beta}_k X_{ki}$$

trong đó  $\hat{Y}_i$  là ước lượng cho  $Y_i$  và phần dư  $e_i = Y_i - \hat{Y}_i$ .

Bài toán bình phương bé nhất trong mô hình hồi quy k biến có nội dung là:

Tìm  $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$  sao cho  $\sum e_i^2 \rightarrow \min$ .

Hay  $\sum e_i^2 = \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_{2i} - \dots - \hat{\beta}_k X_{ki})^2 \rightarrow \min$ .

Như vậy để tìm các ước lượng  $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$  ta cần giải hệ phương trình sau đây:

$$\begin{cases} \sum_{i=1}^n (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_{2i} - \dots - \hat{\beta}_k X_{ki}) = 0 \\ \sum_{i=1}^n X_{2i} (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_{2i} - \dots - \hat{\beta}_k X_{ki}) = 0 \\ \dots \dots \dots \dots \dots \dots \\ \sum_{i=1}^n X_{ki} (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_{2i} - \dots - \hat{\beta}_k X_{ki}) = 0 \end{cases}$$

Với điều kiện số quan sát trong mẫu lớn hơn số hệ số cần ước lượng ( $n > k$ ) và giả thiết không xảy ra đa cộng tuyến hoàn hảo được thoả mãn thì hệ phương trình trên sẽ có nghiệm duy nhất.

**Ví dụ:** xét mô hình hồi quy 2 biến  $Y_i = \beta_1 + \beta_2 X_i + u_i$

Kết quả ước lượng OLS là:

$$\begin{cases} \hat{\beta}_2 = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2} \\ \hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X} \end{cases}$$

trong đó  $x_i = X_i - \bar{X}$ ,  $y_i = Y_i - \bar{Y}$ .

Với mô hình hồi quy bội 3 biến

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$$

thì ta có:

$$\begin{cases} \hat{\beta}_2 = \frac{(\sum x_{2i} y_i)(\sum x_{3i}^2) - (\sum x_{3i} y_i)(\sum x_{2i} x_{3i})}{(\sum x_{2i}^2)(\sum x_{3i}^2) - (\sum x_{2i} x_{3i})^2} \\ \hat{\beta}_3 = \frac{(\sum x_{3i} y_i)(\sum x_{2i}^2) - (\sum x_{2i} y_i)(\sum x_{2i} x_{3i})}{(\sum x_{2i}^2)(\sum x_{3i}^2) - (\sum x_{2i} x_{3i})^2} \\ \hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}_2 - \hat{\beta}_3 \bar{X}_3 \end{cases}$$

## 2.3. Hướng dẫn Eviews

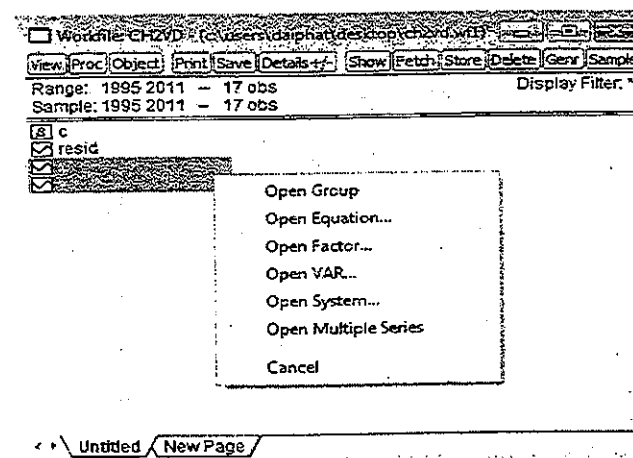
### 2.3.1. Ví dụ

Cho số liệu về tiêu dùng cuối cùng theo giá thực tế (Y) và GDP của Việt Nam tính theo giá thực tế (X) từ năm 1995 – 2011 (đơn vị: tỷ đồng).

Năm	X	Y	Năm	X	Y
1995	228892	187233	2004	715307	511221
1996	272036	225231	2005	839211	584793
1997	313623	250584	2006	974266	675916
1998	361016	283444	2007	1143715	809862
1999	399942	301690	2008	1485038	1091876
2000	441646	321853	2009	1658389	1206819
2001	481295	342607	2010	1980914	1446901
2002	535762	382137	2011	2535008	1794465
2003	613443	445221			

Nguồn: Tổng cục thống kê Việt Nam.

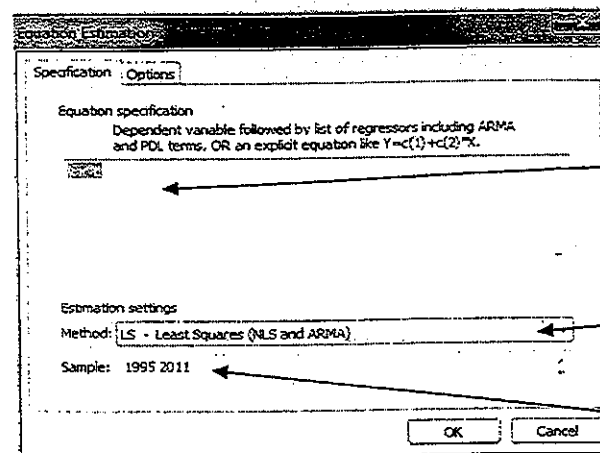
Sau khi nhập số liệu ta đặt tên nó là **ch2vd.wfl** rồi thực hiện ước lượng OLS bằng Eviews như sau:



Chọn hai biến XY rồi nhấp đúp chuột trái ta được hộp thoại → chọn **Open Equation...**

Lưu ý: khi có một danh sách nhiều biến ta có thể vừa nhấn phím Ctrl vừa chọn từng biến.

Sau khi chọn Open Equation ta được hộp thoại sau:



Phân khai báo dạng hàm hồi quy. Khai báo như bên trên là hàm hồi quy có dạng:  
 $Y_i = \beta_1 + \beta_2 X_i + u_i$   
(C chính là hằng số, biến gắn với hệ số chặn)

Phương pháp ước lượng

Phạm vi mẫu muốn ước lượng

Trong cửa sổ khai báo Equation specification thì mặc định hằng số C luôn là biến cuối, chúng ta có thể đặt lại sau biến phụ thuộc tùy ý. Vị trí đầu tiên trong danh sách khai báo biến luôn là biến phụ thuộc, các biến cách nhau một khoảng trống. Trong Eviews cũng không có ký hiệu  $\beta$  nên nó dùng ký hiệu c(1), c(2)

... để thay thế. Ví dụ hai cách khai báo sau là hoàn toàn như nhau cho một dạng hàm hồi quy:

$$Y C \log(X) \text{ (hoặc } Y \log(X) C \text{)}$$

$$Y = c(1) + c(2) * \log(X) \text{ (hoặc } Y = c(1) * \log(X) + c(2) \text{)}$$

Điều là chỉ dạng hàm hồi quy:  $Y_i = \beta_1 + \beta_2 \ln X_i + u_i$   
 ( $Y_i = \beta_1 \ln X_i + \beta_2 + u_i$ )

Trong **Estimation settings** ta có thể chọn các phương pháp ước lượng khác phù hợp với mục đích hồi quy:

LS - Least Squares (NLS and ARMA)

TSLS - Two-Stage Least Squares (TSNLS and ARMA)

GMM - Generalized Method of Moments

ARCH - Autoregressive Conditional Heteroskedasticity

BINARY - Binary Choice (Logit, Probit, Extreme Value)

ORDERED - Ordered Choice

CENSORED - Censored or Truncated Data (including Tobit)

COUNT - Integer Count Data

QREG - Quantile Regression (including LAD)

STEPLS - Stepwise Least Squares

Với phương pháp OLS thì ta chọn LS - Least Squares (NLS and ARMA)

Trong lựa chọn **Sample** ta cũng có thể chọn lại phạm vi mẫu muốn ước lượng bằng cách chỉnh sửa lại thời điểm bắt đầu và thời điểm kết thúc.

Khi lựa chọn xong ta nhấn OK thì được bảng kết quả Eviews sau đây:

Equation: UNTITLED - Workfile: CP12VD - Unrestricted

Dependent Variable: Y  
 Method: Least Squares  
 Date: 04/23/14 Time: 14:47  
 Sample: 1995 2011  
 Included observations: 17

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X	0.709561	0.006771	104.8053	0.0000
C	13616.43	7414.642	1.836495	0.0862

R-squared 0.998536 Mean dependent var 638932.5  
 Adjusted R-squared 0.998545 S.D. dependent var 45885.0  
 S.E. of regression 18150.25 Akaike info criterion 22.56089  
 Sum squared resid 4.94E+09 Schwarz criterion 22.65891  
 Log likelihood -189.7675 Hannan-Quinn criter 22.57098  
 F-statistic 10984.15 Durbin-Watson stat 0.947434  
 Prob(F-statistic) 0.000000

Các lựa chọn liên quan đến tính ma trận hiệp phương sai, đồ thị phân dư, kiểm định khuyết tật

Đặt tên cho Equation này để tiện cho việc lưu trữ về sau

Chỉnh sửa lại dạng hàm hồi quy

Trong mục **[Equation] View** ta có các tùy chọn sau:

Representations

Estimation Output

Actual, Fitted, Residual

ARMA Structure...

Gradients and Derivatives

Covariance Matrix

Coefficient Tests

Residual Tests

Stability Tests

Label

Representation và Estimation Output: Trình bày kết quả

Các vấn đề liên quan đến thông kê mô tả và ma trận hiệp phương sai

Nhóm lệnh kiểm định  
 Coefficient Test: kiểm định hệ số hồi quy  
 Residual Test: kiểm định liên quan đến nhiễu  
 Stability Test: kiểm định tính ổn định của dạng hàm

### 2.3.2. Đọc bảng kết quả Eviews

**Dependent Variable:** Y  
**Method:** Least Squares  
**Date:** Time:  
**Sample:** 1995 2011  
**Included observations:** 17

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X	0.709561	0.006771	104.8053	0.0000
C	13616.43	7414.642	1.836495	0.0862

X	0.709661	0.006771	104.8053	0.0000
C	13616.43	7414.642	1.836425	0.0862

R-squared	0.998636	Mean dependent var	638932.5
Adjusted R-squared	0.998545	S.D. dependent var	475885.0
S.E. of regression	18150.25	Akaike info criterion	22.56089
Sum squared resid	4.94E+09	Schwarz criterion	22.65891
Log likelihood	-189.7675	Hannan-Quinn criter.	22.57063
F-statistic	10984.15	Durbin-Watson stat	0.947434
Prob(F-statistic)	0.000000		

### Ý nghĩa của các thuật ngữ:

Dependent Variable: *biến phụ thuộc*

Method: Least Squares: *phương pháp: bình phương tối thiểu*

Date ... Time ... : ngày giờ tạo file

Sample: 1995 2011 : *mẫu được lấy từ năm 1995 - 2011*

Included observations: *số lượng quan sát n = 17*

Variable: *biến độc lập* Log likelihood : *Giá trị của hàm hợp lý*

Coefficient : *hệ số ước lượng* F-statistic: *F quan sát trong kiểm định về sự phù hợp của hàm hồi quy*

Std. Error : *độ lệch tiêu chuẩn, sai số tiêu chuẩn* Mean dependent var: *giá trị trung bình của biến phụ thuộc*

t-statistic : *T quan sát trong kiểm định ý nghĩa thống kê của hệ số hồi quy* S.D. dependent var: *độ lệch chuẩn của biến phụ thuộc:*

Prob: *giá trị p-value (của kiểm định nào chỉ ứng với giá trị quan sát của kiểm định đó)*

$$SD_y = \sqrt{\frac{\sum (Y_i - \bar{Y})^2}{n-1}} = \sqrt{\frac{TSS}{n-1}}$$

R-squared : *Hệ số xác định* Akaike info criterion : *tiêu chuẩn Akaike*

Adjusted R-squared: *hệ số xác định đã điều chỉnh* Schwarz criterion : *tiêu chuẩn Schwarz*

S.E. of regression : *sai số tiêu chuẩn của đường hồi quy* Hannan-Quinn criter.: *tiêu chuẩn Hannan - Quinn*

Sum squared resid : *tổng của bình phương các phần dư* Durbin-Watson stat : *thống kê Durbin - Watson*

**Lưu ý:** Lệnh xem bảng kết quả Eviews [Equation] View → Estimation Output

Nếu sử dụng cách khai báo thứ nhất:  $Y = C(1) + C(2)*X$  thì kết quả chỉ khác phần thể hiện các hệ số như sau:

$$Y=C(1)+C(2)*X$$

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	13616.43	7414.642	1.836425	0.0862
C(2)	0.709661	0.006771	104.8053	0.0000

Các thông tin khác vẫn giữ nguyên

Bên cạnh cách như trên, còn một số cách sau:

**Cách 2:** [Eviews] Objects → New Object → Equation → OK

**Cách 3:** Chọn X, Y thành cửa sổ [Group] Procs → Make Equation

**Cách 4:** Chọn X và Y, nhấn chuột phải → Open → As Equation

**Cách 5:** [Cửa sổ lệnh] LS Y C X

### 2.3.3. Lưu kết quả Eviews

Sau khi ước lượng được một mô hình hồi quy, ta nên lưu lại để lần sau thực hiện các công việc khác cho tiện lợi. Thực hiện như sau:

**[Equation] Name:** cửa số [Object Name], tên ngầm định cho kết quả ước lượng phương trình hồi quy là Eq01. Nếu chọn OK, thì chủ thể  $\Xi$  Eq01 được tạo ra trong Workfile.

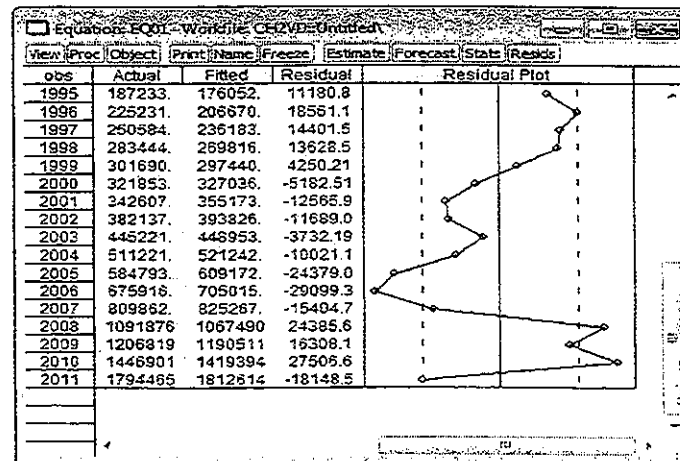
Muốn sao và dán kết quả ước lượng để chèn vào các văn bản, đánh dấu toàn bộ bảng kết quả chi tiết, nhấn chuột phải, chọn Copy, chọn loại giữ nguyên định dạng hoặc không giữ định dạng, rồi tại nơi cần dán chọn Paste; hoặc sử dụng tổ hợp Ctrl+C và Ctrl+V.

### 2.3.4. Xem giá trị thật, giá trị ước lượng và giá trị phần dư

Sau khi ước lượng một mô hình hồi quy, có thể đánh giá kết quả ước lượng thông qua các giá trị phần dư (residuals) và các giá trị ước lượng của biến phụ thuộc (còn gọi là giá trị tương hợp – fitted values).

**[Equation]View→Actual, Fitted, Residual→Actual, Fitted, Residual Table**

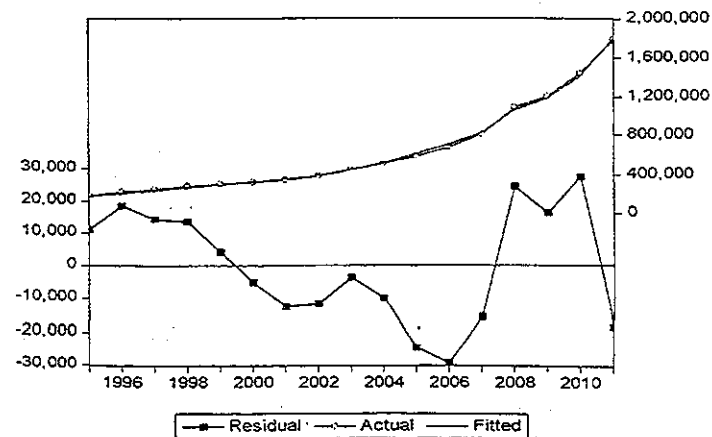
Được bảng giá trị và đồ thị phần dư, Actual là giá trị thực tế bộ số liệu: Y, Fitted là giá trị ước lượng bởi hàm hồi quy  $\hat{Y}_i$ , Residual là giá trị phần dư.



- Quan sát nào có giá trị thực tế và giá trị ước lượng gần nhau nhất, xa nhau nhất? Khi đó giá trị thực tế cao hơn hay thấp hơn giá trị ước lượng?

- Với những quan sát nào thì đường hồi quy mô tả gần đúng nhất sự biến động của biến phụ thuộc Y?

Có thể xem đồ thị của giá trị thực tế, giá trị ước lượng, và phần dư trên cùng hệ tọa độ hoặc đồ thị của riêng phần dư bằng cách chọn Actual, Fitted, Residual Graph. Các đồ thị này có thể lưu lại hoặc cắt dán vào các văn bản.



## 2.4. Bài tập

**Bài 2.1.** Các mô hình hồi quy sau đây có phải là dạng tuyến tính hay không?

- a.  $Y_i = \exp(\beta_1 + \beta_2 X_i + u_i)$       b.  $Y_i = \beta_1 + \beta_2 \ln(X_i) + u_i$   
 c.  $Y_i = 1/[1 + \exp(\beta_1 + \beta_2 X_i + u_i)]$       d.  $Y_i = \beta_1^2 + \beta_2 X_i + u_i$   
 e.  $Y_i = X_i / [\beta_1 + \beta_2 X_i]$       f.  $Y_i = 1/[\beta_1 + \beta_2 X_i]$

Hãy biến đổi mô hình không phải tuyến tính về dạng tuyến tính nếu được?

**Bài 2.2.** Cho bộ số liệu sau đây với Y – lượng bia tiêu thụ; X – thu nhập.

Đv: 100.000đ

STT	Y	X	X <sup>2</sup>	XY	$\hat{Y}_i$	$e_i = Y_i - \hat{Y}_i$
1	10	30	900	300	10.0697	0.004858
2	10	32	1024	320	11.2129	1.471126
3	12	35	1225	420	12.9277	0.860627
4	13	35	1225	455	12.9277	0.005227
5	15	38	1444	570	14.6425	0.127806
6	18	40	1600	720	15.7857	4.903124
7	18	42	1764	756	16.9289	1.147255
8	19	45	2025	855	18.6437	0.12695
9	20	48	2304	960	20.3585	0.128522
10	20	50	2500	1000	21.5017	2.255103
<b>Tổng</b>	<b>155</b>	<b>395</b>	<b>16011</b>	<b>6356</b>		<b>11.0306</b>

Cho biết mô hình hồi quy là :  $Y_i = \beta_1 + \beta_2 X_i + u_i$

- Hãy xác định các ước lượng của tham số hồi quy.
- Kết quả ước lượng có phù hợp với lý thuyết kinh tế không?
- Tính hệ số xác định và cho biết ý nghĩa của nó?

**Bài 2.3.** Có số liệu về giảng viên một trường đại học như sau: tiền lương Y (triệu đồng/tháng), số năm công tác X (năm). Cho mức ý nghĩa  $\alpha = 5\%$

Y	3	2.7	4	4.5	4	4.2	5	6	7	6
X	1	2	3	4	5	6	7	8	9	10

Xét mô hình:  $Y_i = \beta_1 + \beta_2 X_i + u_i$

- Hãy xác định hàm SRF của (1). Nêu ý nghĩa của các hệ số ước lượng?
- Tính hệ số xác định và cho biết ý nghĩa của nó?

**Bài 2.4.** Để tìm hiểu xem giá bán thịt heo (đơn vị: nghìn đồng/kg) và giá bán thịt gà (đơn vị: nghìn đồng/kg) có ảnh hưởng lên mức tiêu thụ thịt heo không (kg/ngày), ta khảo sát mẫu số liệu sau đây từ quý 1/2005 đến quý 4/2007 về mức tiêu thụ thịt heo bình quân một ngày tại một siêu thị:

Quan sát	Giá thịt heo	Giá thịt gà	Thịt heo tiêu thụ
2005Q1	45	85	1200
2005Q2	49	81	1176
2005Q3	54	76	1152
2005Q4	58	73	1087
2006Q1	63	70	1045
2006Q2	69	72	1023
2006Q3	72	68	985
2006Q4	73	63	942
2007Q1	76	59	915
2007Q2	80	55	845
2007Q3	83	54	810
2007Q4	86	52	759

1. Ước lượng hàm hồi quy dạng: Tuyến tính; Lin – Log; Log – Log?
2. Cho biết ý nghĩa của các hệ số góc trong mỗi mô hình?
3. Chọn mô hình phù hợp nhất từ các mô hình trên, giải thích tại sao chọn?
4. Nhận xét ý nghĩa kinh tế các tham số hồi quy từ mô hình được chọn ở câu 3? Chúng có phù hợp với lý thuyết kinh tế không?

**Bài 2.5.** Cho các số liệu về PCE – PCE services (chi tiêu tiêu dùng cá nhân cho dịch vụ) (tỷ USD) GDP – Gross domestic product (tổng sản phẩm quốc dân) (tỷ USD). Tiến hành khảo sát số liệu kinh tế Úc từ quý I năm 1983 đến quý I năm 2008.

Dependent Variable: PCE  
Method: Least Squares  
Sample: 1983Q1 2008Q1  
Included observations: 101

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP	0.407484	0.001967	207.1676	0.0000
C	-41.46932	16.89039	-2.455202	0.0158

R-squared		Mean dependent var	3376.426
Adjusted R-squared	0.997675	S.D. dependent var	754.2616
S.E. of regression	36.36636	Akaike info criterion	10.04477
Sum squared resid	130928.7	Schwarz criterion	10.09655
Log likelihood	-505.2608	Hannan-Quinn criter.	10.06573
F-statistic	42918.42	Durbin-Watson stat	0.248554
Prob(F-statistic)	0.000000		

1. Từ bảng kết quả Eviews hãy viết hàm hồi quy tổng thể và hàm hồi quy mẫu.
2. Tính  $R^2$  từ các đại lượng đã biết?
3. Hãy cho biết tổng mức PCE của 101 quý là bao nhiêu?
4. Hãy tính mức GDP trung bình từ quý I năm 1983 tới quý I năm 2008.

**Bài 2.6.** File **ch2.6** là số liệu về GDP của Việt Nam theo giá so sánh năm 1994 giai đoạn 1990-2012 (đơn vị: tỉ VNĐ).

1. Vẽ biểu đồ GDP theo thời gian t và cho nhận xét?
2. Xác định mức độ gia tăng bình quân tính theo số tuyệt đối của GDP qua mỗi năm?
3. Xác định tốc độ tăng trưởng bình quân của GDP trong giai đoạn này?

**Bài 2.7.** File **ch2.7** là số liệu của Singapore về chỉ số giảm phát GDP đối với hàng nội địa (Y) và chỉ số giảm phát GDP đối với hàng nhập khẩu (X) trong giai đoạn 1968-1982. Chỉ số giảm phát GDP thường được sử dụng làm một chỉ số cho lạm phát thay cho chỉ số giá hàng tiêu dùng CPI. Singapore là một nền kinh tế mở với quy mô nhỏ, phụ thuộc nhiều vào ngoại thương.

Để nghiên cứu mối quan hệ giữa giá nội địa và giá thế giới, bạn được cho biết các mô hình sau:

$$Y_i = \beta_1 + \beta_2 X_i + u_i \quad (1)$$

$$Y_i = \beta_2 X_i + v_i \quad (2)$$

1. Vẽ đồ thị của các chỉ số giảm phát theo thời gian và cho nhận xét?
2. Hãy ước lượng hai mô hình trên và xác định xem mô hình nào thích hợp hơn với lý thuyết kinh tế và sự biến động?



**Bài 2.8.** Cho các số liệu về giá trị nhập khẩu (IMPORTS: triệu USD), GDP (tỉ VNĐ) và chỉ số giá tiêu dùng (CPI, 1994 = 100) trong thời kỳ 1995-2013 sau đây:

Năm	IMPORTS	CPI	GDP
1995	8155.4	117.04	195567
1996	11143.6	127.22	213833
1997	11592.3	135.61	231264
1998	11499.6	147.6	244595
1999	11742.1	156.06	256269
2000	15636.5	161.38	273666
2001	16218	164.53	292535
2002	19745.6	171.03	313247
2003	25255.8	182.44	336242
2004	31968.8	197.36	362435
2005	36761.1	213.52	393031
2006	44891.1	229.04	425374
2007	62764.7	247.91	461344
2008	80713.8	301.68	489833
2009	69948.8	318.95	515909
2010	84838.6	359.12	551609
2011	106749.8	434.02	584073
2012	113790	463.58	615029
2013	132130	491.58	648241

1. Vẽ đồ thị của các biến số và cho nhận xét về sự biến động của chúng theo thời gian?

2. Hãy ước lượng các mô hình tuyến tính sau đây:

$$IMPORT = \beta_1 + \beta_2 GDP + \beta_3 CPI + u \quad (1)$$

$$\ln(IMPORT) = \beta_1 + \beta_2 \ln(GDP) + \beta_3 \ln(CPI) + u \quad (2)$$

$$IMPORT = \beta_1 + \beta_2 \ln(GDP) + \beta_3 \ln(CPI) + u \quad (3)$$

$$\ln(IMPORT) = \beta_1 + \beta_2 GDP + \beta_3 CPI + u \quad (4)$$

3. Giải thích ý nghĩa của các hệ số góc nhận được trong mỗi một mô hình?

4. Mô hình nào là phù hợp nhất cho mối quan hệ giữa Import với GDP, CPI?

## Chương 3: SUY DIỄN THỐNG KÊ

### 3.1. Các khoảng tin cậy và kiểm định giả thiết

#### 3.1.1. Khoảng tin cậy

##### Khoảng tin cậy cho các hệ số hồi quy

Dùng để đánh giá tác động của từng biến độc lập lên biến phụ thuộc.

- Khoảng tin cậy đối xứng (hai phía):

$$(\hat{\beta}_j - se(\hat{\beta}_j).T_{\alpha/2}(n-k); \hat{\beta}_j + se(\hat{\beta}_j).T_{\alpha/2}(n-k))$$

Trong đó:  $se(\hat{\beta}_j)$  là độ lệch chuẩn hoặc sai số tiêu chuẩn của ước lượng OLS,  $n$  là số quan sát (kích cỡ mẫu),  $k$  là số hệ số hồi quy của mô hình,  $T_{\alpha/2}(n-k)$  là giá trị tới hạn mức  $\alpha/2$  của phân phối Student với bậc tự do là  $(n-k)$ .

- Khoảng tin cậy bên phải (ước lượng giá trị tối thiểu):

$$\beta_j > \hat{\beta}_j - se(\hat{\beta}_j).T_{\alpha}(n-k)$$

- Khoảng tin cậy bên trái (ước lượng giá trị tối đa):

$$\beta_j < \hat{\beta}_j + se(\hat{\beta}_j).T_{\alpha}(n-k)$$

Khoảng tin cậy bên phải để ước lượng giá trị tối thiểu của  $\beta_j$ , khoảng tin cậy bên trái để ước lượng giá trị tối đa của  $\beta_j$ .

##### Khoảng tin cậy cho biểu thức của hai hệ số hồi quy

Khi cả hai biến độc lập cùng thay đổi một lượng nào đó, thì biến phụ thuộc sẽ biến thiên trong khoảng nào?

Khi biến độc lập  $X_i$  thay đổi  $a$  đơn vị, biến độc lập  $X_j$

cũng thay đổi  $b$  đơn vị thì trung bình biến phụ thuộc  $Y$  sẽ thay đổi  $a\hat{\beta}_i + b\hat{\beta}_j$  đơn vị.

- Khoảng tin cậy hai phía (đối xứng):

$$((a\hat{\beta}_i + b\hat{\beta}_j) - se(a\hat{\beta}_i + b\hat{\beta}_j).T_{\alpha/2}(n-k); (a\hat{\beta}_i + b\hat{\beta}_j) + se(a\hat{\beta}_i + b\hat{\beta}_j).T_{\alpha/2}(n-k))$$

Với

$$se(a\hat{\beta}_i + b\hat{\beta}_j) = \sqrt{a^2 Var(\hat{\beta}_i) + b^2 Var(\hat{\beta}_j) + 2ab.cov(\hat{\beta}_i, \hat{\beta}_j)}$$

- Khoảng tin cậy bên phải (ước lượng giá trị tối thiểu):

$$a\hat{\beta}_i + b\hat{\beta}_j > (a\hat{\beta}_i + b\hat{\beta}_j) - se(a\hat{\beta}_i + b\hat{\beta}_j).T_{\alpha}(n-k)$$

- Khoảng tin cậy bên trái (ước lượng giá trị tối đa):

$$a\hat{\beta}_i + b\hat{\beta}_j < (a\hat{\beta}_i + b\hat{\beta}_j) + se(a\hat{\beta}_i + b\hat{\beta}_j).T_{\alpha}(n-k)$$

##### Khoảng tin cậy cho phương sai của sai số ngẫu nhiên

Phương sai của sai số ngẫu nhiên là  $\sigma^2$  và ước lượng điểm của nó là:

$$\hat{\sigma}^2 = \frac{RSS}{n-k}$$

Ta cũng có thể xác định các khoảng biến thiên cho nó như sau:

- Khoảng tin cậy hai phía:

$$\frac{\hat{\sigma}^2(n-k)}{\chi_{\alpha/2}^2(n-k)} < \sigma^2 < \frac{\hat{\sigma}^2(n-k)}{\chi_{1-\alpha/2}^2(n-k)}$$

- Khoảng tin cậy bên phải:  $\frac{\hat{\sigma}^2(n-k)}{\chi_{\alpha}^2(n-k)} < \sigma^2$

- Khoảng tin cậy bên trái:  $\sigma^2 < \frac{\hat{\sigma}^2(n-k)}{\chi_{1-\alpha}^2(n-k)}$

### Ý nghĩa của các khoảng tin cậy

Với độ tin cậy 95%, khoảng tin cậy cho hệ số  $\beta_j$  được hiểu như sau: nếu lấy các mẫu nhiều lần một cách ngẫu nhiên từ tổng thể hồi quy, sẽ có khoảng 95% số khoảng tin cậy được xây dựng từ các mẫu này chứa giá trị  $\beta_j$ .

Trong thực tế hồi quy, ta chỉ lấy một mẫu duy nhất và một khoảng tin cậy cụ thể cho  $\beta_j$  nên ta hi vọng rằng khoảng tin cậy này nằm trong số 95% đó.

Khi độ tin cậy  $(1 - \alpha)$  càng cao thì độ dài của khoảng tin cậy càng rộng, nghĩa là độ chính xác càng giảm và ít có giá trị thông tin. Hiện nay mức tin cậy 95% là mức chuẩn, thường được sử dụng nhất.

### 3.1.2. Các kiểm định cơ bản cho mô hình hồi quy

#### 3.1.2.1. Kiểm định về giá trị của hệ số hồi quy

Phương pháp kiểm định giá trị tới hạn cho các cặp giả thiết

Loại kiểm định	Cặp giả thiết	Tiêu chuẩn kiểm định	Miền bác bỏ $H_0$
Hai phía	$H_0: \beta_j = a$ $H_1: \beta_j \neq a$	$T_{qs} = \frac{\hat{\beta}_j - a}{se(\hat{\beta}_j)}$	$ T_{qs}  > t_{\alpha/2}^{(n-k)}$
Bên phải	$H_0: \beta_j \leq a$ $H_1: \beta_j > a$		$T_{qs} > t_{\alpha}^{(n-k)}$
Bên trái	$H_0: \beta_j \geq a$ $H_1: \beta_j < a$		$T_{qs} < -t_{\alpha}^{(n-k)}$

**Lưu ý:** giả thiết  $H_0$  trong các kiểm định giả thiết thống kê về hệ số hồi quy luôn chứa dấu "=", yêu cầu này liên quan đến kỹ thuật kiểm định và nếu lẫn lộn giữa  $H_0$  và  $H_1$  thì sẽ dẫn đến các kết luận sai lầm.

#### 3.1.2.2. Kiểm định về sự ràng buộc giữa các hệ số hồi quy

Chẳng hạn, ta cần so sánh tác động của hai biến độc lập lên biến phụ thuộc hay cần kiểm tra nhận định về biến phụ thuộc khi hai biến độc lập cùng tăng, cùng giảm, ... khi đó ta cần kiểm định cho hai hệ số hồi quy.

Loại kiểm định	Cặp giả thiết	Tiêu chuẩn kiểm định	Miền bác bỏ $H_0$
Hai phía	$H_0: a\beta_i + b\beta_j = c$ $H_1: a\beta_i + b\beta_j \neq c$	$T_{qs} = \frac{(a\hat{\beta}_i + b\hat{\beta}_j) - c}{se(a\hat{\beta}_i + b\hat{\beta}_j)}$	$ T_{qs}  > T_{\alpha/2}^{(n-k)}$
Bên phải	$H_0: a\beta_i + b\beta_j \leq c$ $H_1: a\beta_i + b\beta_j > c$		$T_{qs} > T_{\alpha}^{(n-k)}$
Bên trái	$H_0: a\beta_i + b\beta_j \geq c$ $H_1: a\beta_i + b\beta_j < c$		$T_{qs} < -T_{\alpha}^{(n-k)}$

Trong đó  $a, b, c$  là các hằng số cho trước, tùy vào yêu cầu kiểm định.

Ghi chú: Khi thực hiện các kiểm định trên, ngoài phương pháp sử dụng giá trị tới hạn, ta có thể đưa ra kết luận dựa trên giá trị xác suất P (hay P-value). Giá trị xác suất P được định nghĩa là mức ý nghĩa nhỏ nhất mà giả thiết  $H_0$  bị bác bỏ tương ứng với giá trị quan sát của thống kê kiểm định này.

Trong các bảng kết quả ước lượng từ Eviews, giá trị trong cột Prob. Và dòng thứ j là giá trị xác suất P trong bài toán kiểm định cặp giả thiết:

$$H_0: \beta_j = 0; H_1: \beta_j \neq 0$$

Từ cột này, ta có thể kết luận biến  $X_j$  có tác động đến biến phụ thuộc hay không. Trong thực hành, mức ý nghĩa  $\alpha$  thường được chọn là 5% và khi đó với giá trị P-value  $< 0,05$  thì giả thiết  $H_0$  bị bác bỏ. Trường hợp  $H_0: \beta_j \leq 0$  hoặc  $H_0: \beta_j \geq 0$  thì so sánh P-value với  $\alpha/2$ .

Đối với các kiểm định khác sau này, cách kiểm định dựa

trên giá trị xác suất P được thực hiện tương tự.

### 3.1.2.3. Kiểm định thu hẹp hàm hồi quy

Xét mô hình ban đầu (gọi là mô hình nhiều biến,  $nb$ ):

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_m X_m + \beta_{m+1} X_{m+1} + \dots + \beta_{k-1} X_{k-1} + u$$

Ta nghi ngờ  $m$  biến độc lập  $X_1, \dots, X_m$  không tác động tới biến phụ thuộc  $Y$ . Vì vậy ta loại bỏ  $m$  biến này ra khỏi mô hình ban đầu. Mô hình mới (gọi là mô hình ít biến,  $ib$ ):

$$Y = \beta_0 + \beta_{m+1} X_{m+1} + \dots + \beta_{k-1} X_{k-1} + u$$

Muốn kiểm tra xem việc loại bỏ  $m$  biến này có hợp lý hay không ta thực hiện kiểm định cặp giả thiết sau đây:

$$H_0: \beta_1 = \dots = \beta_m = 0 \quad \Leftrightarrow \quad H_0: \text{mô hình ít biến đúng}$$

$$H_1: \exists \beta_j \neq 0, j = 1, m \quad H_1: \text{mô hình nhiều biến đúng}$$

Tiêu chuẩn kiểm định

$$F_{qs} = \frac{(R_{nb}^2 - R_{ib}^2) / m}{(1 - R_{nb}^2) / (n - k)}$$

Trong đó:  $m$  là số biến chênh lệch giữa hai mô hình,  $k$  là số tham số của mô hình nhiều biến.

Nếu  $F_{qs} > F_{\alpha}(m; n - k)$  thì bác bỏ giả thiết  $H_0$ .

Kiểm định thu hẹp hồi quy còn dùng trong trường hợp mở rộng hàm hồi quy. Trong cả hai kiểm định thì cặp giả thiết, tiêu chuẩn kiểm định và miền bác bỏ hoàn toàn như nhau.

### 3.1.2.4. Kiểm định về sự phù hợp của hàm hồi quy

Ta cần thực hiện kiểm định cặp giả thiết sau đây:

$H_0$ : mô hình hồi quy đưa ra không phù hợp.

$H_1$ : mô hình hồi quy đưa ra là phù hợp.

Tiêu chuẩn kiểm định:

$$F_{qs} = \frac{R^2 / (k - 1)}{(1 - R^2) / (n - k)}$$

Nếu  $F_{qs} > F_{\alpha}(k - 1; n - k)$  thì bác bỏ  $H_0$ : hàm hồi quy là phù hợp.

### 3.1.2.5. Kiểm định về phương sai nhiễu

Loại kiểm định	Cặp giả thiết	Tiêu chuẩn kiểm định	Miền bác bỏ $H_0$
Hai phía	$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 \neq \sigma_0^2$	$\chi_{qs}^2 = \frac{(n - k) \hat{\sigma}^2}{\sigma_0^2}$	$\chi_{qs}^2 > \chi_{\frac{\alpha}{2}}^2(n - k)$ Hoặc $\chi_{qs}^2 < \chi_{1 - \frac{\alpha}{2}}^2(n - k)$
Bên phải	$H_0: \sigma^2 \leq \sigma_0^2$ $H_1: \sigma^2 > \sigma_0^2$		$\chi_{qs}^2 > \chi_{\alpha}^2(n - k)$
Bên trái	$H_0: \sigma^2 \geq \sigma_0^2$ $H_1: \sigma^2 < \sigma_0^2$		$\chi_{qs}^2 < \chi_{1 - \alpha}^2(n - k)$

### 3.1.3. Dự báo giá trị của biến phụ thuộc

Xét mô hình hồi quy bội  $k$  biến

$$Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i$$

Cho

$$X^0 = \begin{bmatrix} 1 \\ X_2^0 \\ \dots \\ X_k^0 \end{bmatrix}$$

Dự báo giá trị trung bình có điều kiện của biến phụ thuộc  $Y$  với  $X = X^0$ , tức là dự báo:

$$E(Y|X^0) = \beta_1 + \beta_2 X_2^0 + \dots + \beta_k X_k^0$$

Với độ tin cậy  $1 - \alpha$ , dự báo khoảng của  $E(Y|X^0)$  là:

$$(\hat{Y}_0 - se(\hat{Y}_0) \cdot T_{\alpha/2}(n-k) < E(Y|X^0) < \hat{Y}_0 + se(\hat{Y}_0) \cdot T_{\alpha/2}(n-k))$$

Trong đó,

+  $\hat{Y}_0$  là dự báo điểm (ước lượng điểm) của  $E(Y|X^0)$  xác định bởi:

$$\hat{Y}_0 = \hat{\beta}_1 + \hat{\beta}_2 X_2^0 + \dots + \hat{\beta}_k X_k^0$$

$$+ se(\hat{Y}_0) = \sqrt{Var(\hat{Y}_0)} = \sqrt{\hat{\sigma}^2 (X^0)^T (X^T X)^{-1} X^0}$$

Dự báo khoảng cho giá trị riêng biệt của biến phụ thuộc  $Y$  khi  $X = X^0$ , tức là tìm khoảng tin cậy cho  $Y_0$ :

$$(\hat{Y}_0 - se(Y_0 - \hat{Y}_0) \cdot T_{\alpha/2}(n-k) < Y_0 < \hat{Y}_0 + se(Y_0 - \hat{Y}_0) \cdot T_{\alpha/2}(n-k))$$

$$\text{Trong đó, } se(Y_0 - \hat{Y}_0) = \sqrt{Var(Y_0 - \hat{Y}_0)} = \sqrt{Var(\hat{Y}_0) + \hat{\sigma}^2}$$

Với mô hình hồi quy 2 biến, ta có công thức tính:

$$Var(\hat{Y}_0) = \hat{\sigma}^2 \left[ \frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum_{i=1}^n x_i^2} \right]$$

### 3.2. Hướng dẫn Eviews

#### 3.2.1. Ví dụ 1

Lấy lại ví dụ **CH2VD.Wfl** ta thực hiện một vài kiểm định sau:

Xét mô hình hồi quy:

$$Y_i = \beta_1 + \beta_2 X_i + u_i$$

Ta cần kiểm định cặp giả thiết sau đây:

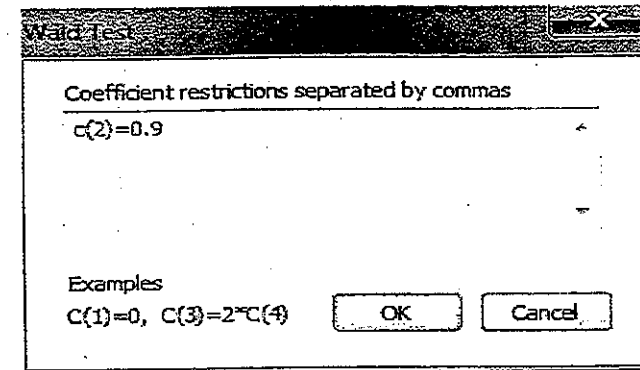
$$H_0 : \beta_2 = 0,9$$

$$H_1 : \beta_2 \neq 0,9$$

Trong bảng kết quả **Equation** ta có các tùy chọn ước lượng và kiểm định như sau:

### [Equation] Views → Coefficient Tests → Wald – Coefficient Restrictions

Ta được hộp thoại sau Wald Test.



Trong Eviews không có ký hiệu beta nên quy ước đánh số các hệ số  $C(1), C(2), \dots$  theo thứ tự khai báo ( $Y = C(1) + C(2)*X$ ) của phương trình hồi quy ban đầu hệ số  $\beta_2$  chính là  $c(2)$ , do đó ta gõ vào ô Coefficient restrictions separated by commas như sau:

$C(2) = 0.9$  để thực hiện cho kiểm định cặp giả thiết trên.

Kết quả kiểm định:

Wald Test:  
Equation: EQ01

Test Statistic	Value	df	Probability
F-statistic	790.1658	(1, 15)	0.0000
Chi-square	790.1658	1	0.0000

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
-0.9 + C(2)	-0.190339	0.006771

Restrictions are linear in coefficients.

Eviews thực hiện kiểm định ràng buộc về các hệ số hồi quy bằng kiểm định Wald, sử dụng thống kê F và thống kê Chi-square.

$$\text{Mô hình ban đầu: } Y_i = \beta_1 + \beta_2 X_i + u_i \quad (\text{nb})$$

$$\text{Mô hình ràng buộc: } Y_i - 0,9X_i = \beta_1 + u_i \quad (\text{ib})$$

$$F_{qs} = \frac{(R_{nb}^2 - R_{ib}^2) / m}{(1 - R_{nb}^2) / (n - k)}$$

Theo kết quả:  $F_{qs} = 790,166 > F_{0,05}(1,15) = 4,543$  nên ta bác bỏ  $H_0$  thừa nhận  $H_1$ . Tức là hệ số  $\beta_2 \neq 0,9$ .

### 3.2.2. Ví dụ 2

Cho các biến số sau:

DT: doanh thu của một bộ phim (đv : tỷ đồng)

CPQC: chi phí quảng cáo (đv : tỷ đồng)

CPSX: chi phí sản xuất (đv : tỷ đồng)

Tập số liệu CH3VD.wfl

Kết quả ước lượng được như sau :

Dependent Variable: DT  
Method: Least Squares  
Sample: 1 20  
Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CPSX	3.715356	0.554553	6.699729	0.0000
CPQC	2.200060	0.387922	5.671394	0.0000
C	5.410282	3.599845	1.502921	0.1512
R-squared	0.914038	Mean dependent var		46.05000
Adjusted R-squared	0.903925	S.D. dependent var		18.55710
S.E. of regression	5.751944	Akaike info criterion		6.474434
Sum squared resid	562.4427	Schwarz criterion		6.623794
Log likelihood	-61.74434	Hannan-Quinn criter.		6.503591
F-statistic	90.38132	Durbin-Watson stat		1.991281
Prob(F-statistic)	0.000000			

Thực hiện lệnh sau để tìm ma trận hiệp phương sai của hệ số hồi quy :

[Equation] View → Covariance Matrix

	CPSX	CPQC	C
CPSX	0.307529	-0.123886	-1.543208
CPQC	-0.123886	0.150484	0.023236
C	-1.543208	0.023236	12.95888

1. Viết hàm hồi quy tổng thể và hàm hồi quy mẫu?

Hàm hồi quy tổng thể:

$$DT_i = \beta_1 + \beta_2 CPSX_i + \beta_3 CPQC_i + u_i$$

Hàm hồi quy mẫu:

$$\widehat{DT}_i = \hat{\beta}_1 + \hat{\beta}_2 CPSX_i + \hat{\beta}_3 CPQC_i$$

$$\widehat{DT}_i = 5,41 + 3,72CPSX_i + 2,20CPQC_i$$

2. Khi chi phí sản xuất không đổi, nếu chi phí quảng cáo tăng 1 tỷ thì doanh thu tăng trong khoảng nào?

Theo yêu cầu ta phải tìm khoảng tin cậy đối xứng cho hệ số của CPQC, tức là hệ số  $\beta_3$ . Khoảng tin cậy đối xứng là:

$$\left( \hat{\beta}_3 - se(\hat{\beta}_3) \cdot T_{\alpha/2}(n - k); \hat{\beta}_3 + se(\hat{\beta}_3) \cdot T_{\alpha/2}(n - k) \right)$$

Với  $T_{\alpha/2}(n - k) = T_{0,025}(17) = 2,1098$  và thay các giá trị từ bảng kết quả Eviews ta được: (1,3816; 3,0185).

Vậy khi chi phí sản xuất không đổi, nếu chi phí quảng cáo tăng 1 tỷ thì doanh thu tăng trong khoảng (1,3816; 3,0185) tỷ đồng.

Ta có thể thực hiện tính toán trực tiếp khoảng tin cậy trên Eviews theo các bước như sau:

Hồi quy mô hình và lưu kết quả hồi quy, giả sử với tên gọi là eq01.

Trên cửa sổ lệnh, ta lần lượt thực hiện các lệnh sau:

**Coef(2) hscpqc** (tên biến này ta do ta tự đặt, 2 là số giá trị ta cần tính cho biến này, ở đây ta tính cận dưới và cận trên của khoảng tin cậy)

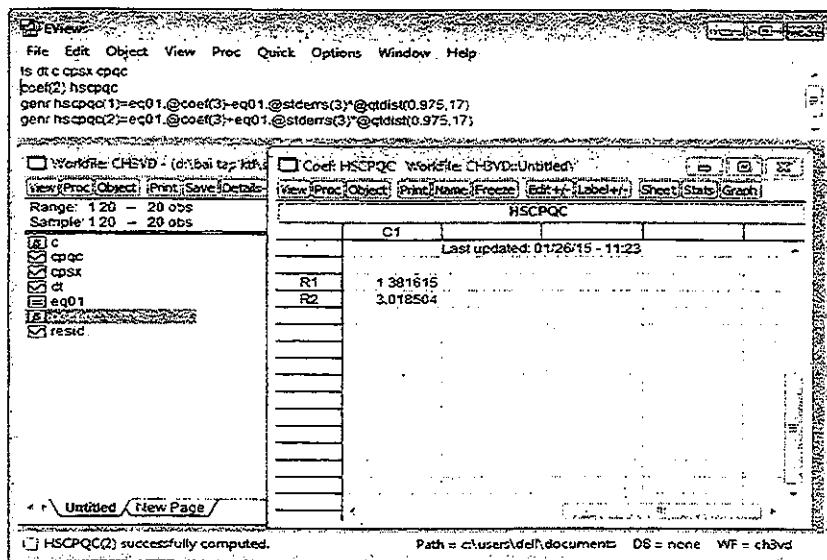
**hscpqc(1)=eq01.@coef(3)-eq01.@stderrs(3)\*@qtdist(0.975,17)**

**hscpqc(2)=eq01.@coef(3)+eq01.@stderrs(3)\*@qtdist(0.975,17)**

@qtdist(0.975,17) là lệnh tìm giá trị tới hạn Student  $T_{0,025}(17)$ .

Số 3 là thứ tự hệ số mà ta cần tính trong bảng hồi quy.

Để xem kết quả khoảng tin cậy ta nhấp đúp vào biểu tượng hscpqc trên cửa sổ Workfile.



3. Phải chăng yếu tố chi phí quảng cáo không ảnh hưởng đến sự biến động về doanh thu của bộ phim?

Theo yêu cầu đề bài ta kiểm định cặp giả thuyết sau:

$$H_0 : \beta_3 = 0$$

$$H_1 : \beta_3 \neq 0$$

Ta có:  $T_{qs} = 5,671$

Mà  $T_{\alpha/2}(n-k) = T_{0,025}(20-3) = 2,11$

Ta thấy:  $|T_{qs}| = 5,671 > 2,11$  nên bác bỏ  $H_0$ , thừa nhận  $H_1$ .

Dựa vào kết quả kiểm định ta kết luận rằng chi phí quảng cáo thực sự có ảnh hưởng đến doanh thu của bộ phim.

Ta cũng có thể thực hiện kiểm định cặp giả thiết trên bằng giá trị p\_value. Theo bảng kết quả Eviews ta có giá trị p\_value tương ứng với giá trị  $T_{qs}$  của biến CPQC bằng 0,0000 (Prob). Giá trị này nhỏ hơn mức ý nghĩa 0,05 (5%) của kiểm định nên ta bác bỏ  $H_0$  và thừa nhận  $H_1$ .

4. Khi CPSX tăng thêm 1 tỷ thì doanh thu tăng thêm tối đa là bao nhiêu?

Theo yêu cầu của đề bài, ta tìm khoảng tin cậy tối đa cho hệ số góc gắn với CPSX, tức là hệ số góc  $\beta_2$ . Khoảng tin cậy này là:

$$(-\infty; \hat{\beta}_2 + se(\hat{\beta}_2) \cdot T_{\alpha}(n-k)) \quad T_{0,05}(17) = 1,74$$

Thay các dữ liệu từ bảng Eviews vào ta được:  $(-\infty; 4,6807)$

Giá trị tối đa ở đây là 4,6807. Vậy có thể kết luận, khi tăng CPSX thêm 1 tỷ thì doanh thu của phim tăng tối đa khoảng 4,6807 tỷ.

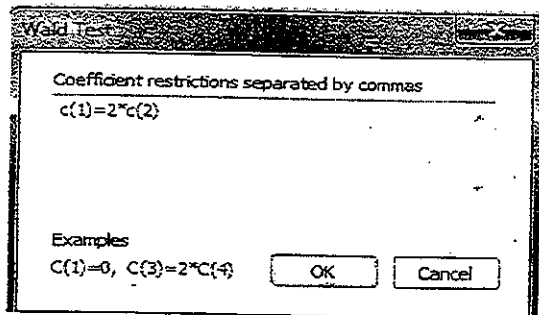
5. Có phải CPSX tác động mạnh gấp đôi lên DT so với CPQC?

CPSX tác động tới sự biến động của doanh thu thông qua giá trị  $|\beta_2|$  còn CPQC là  $|\beta_3|$ . Để biết có phải tác động mạnh gấp đôi hay không ta thực hiện kiểm định sau:

$$\begin{cases} H_0: |\beta_2| = 2|\beta_3| \\ H_1: |\beta_2| \neq 2|\beta_3| \end{cases} \rightarrow \begin{cases} H_0: \beta_2 = 2\beta_3 \\ H_1: \beta_2 \neq 2\beta_3 \end{cases}$$

Để thực hiện kiểm định này ta có thể sử dụng Wald test trong Eviews như sau:

[Equation] Views → Coefficient Tests → Wald – Coefficient Restrictions



Do lệnh ước lượng là:

Estimation Equation:

$$DT = C(1)*CPSX + C(2)*CPQC + C(3)$$

Kết quả kiểm định:

Wald Test:  
Equation: EQ01

Test Statistic	Value	df	Probability
F-statistic	0.333736	(1, 17)	0.5710
Chi-square	0.333736	1	0.5635

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(1) - 2*C(2)	-0.684764	1.185330

Restrictions are linear in coefficients.

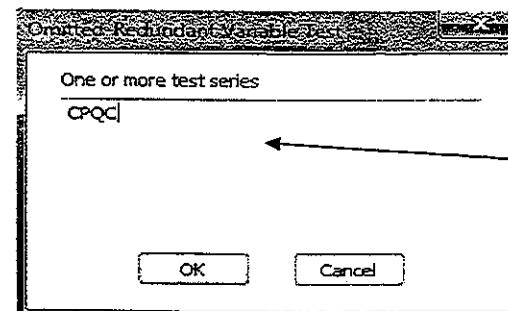
Kết quả ta thấy  $F_{qs} = 0,334 < F_{0,05}(1,17) = 4,451$  vậy chưa đủ cơ sở bác bỏ  $H_0$ . Có thể xem như tác động của CPSX lên DT mạnh gấp đôi CPQC.

6. Có nên bỏ biến CPQC ra khỏi mô hình hay không?

Để trả lời cho câu hỏi này ta có thể thực hiện kiểm định Wald Test hoặc kiểm định ý nghĩa thống kê của hệ số góc. Tuy nhiên trong Eviews có chứa sẵn kiểm định này. Thực hiện lệnh này như sau:

[Equation] Views → Coefficient Tests → Redundant Variables–Likelihood Ratio...

Hộp thoại Omitted-Redundant Variable Test như sau:



Khai báo biến muốn bỏ là CPQC rồi nhấn OK

Đây là một nội dung trong kiểm định sai dạng mô hình (hay kiểm định thừa biến). Kiểm định này cho phép ta kiểm định xem một hoặc một nhóm biến đưa vào mô hình có ý nghĩa thống kê hay không. Nói cách khác, đây là kiểm định xem các hệ số của một nhóm biến đưa vào mô hình có đồng thời bằng 0 hay không để quyết định có nên loại chúng ra khỏi mô hình hay không. Các điều kiện áp dụng kiểm định này như sau:

- Số quan sát trong hai mô hình phải bằng nhau.
- Áp dụng cho mọi phương pháp ước lượng miễn là phương trình hồi quy được xác định bằng cách liệt kê các biến chứ không phải bằng công thức.



Kết quả như sau:

Redundant Variables: CPQC

F-statistic	32.16471	Prob. F(1,17)	0.0000
Log likelihood ratio	21.23926	Prob. Chi-Square(1)	0.0000

Test Equation:

Dependent Variable: DT  
 Method: Least Squares  
 Date: 04/23/14 Time: 23:33  
 Sample: 1 20  
 Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CPSX	5.526557	0.749271	7.375914	0.0000
C	5.070580	5.948592	0.852400	0.4052
R-squared	0.751395	Mean dependent var		46.05000
Adjusted R-squared	0.737584	S.D. dependent var		18.55710
S.E. of regression	9.506161	Akaike info criterion		7.436397
Sum squared resid	1626.608	Schwarz criterion		7.535970
Log likelihood	-72.36397	Hannan-Quinn criter.		7.455835
F-statistic	54.40411	Durbin-Watson stat		1.936532
Prob(F-statistic)	0.000001			

Cặp giả thiết cần kiểm định:

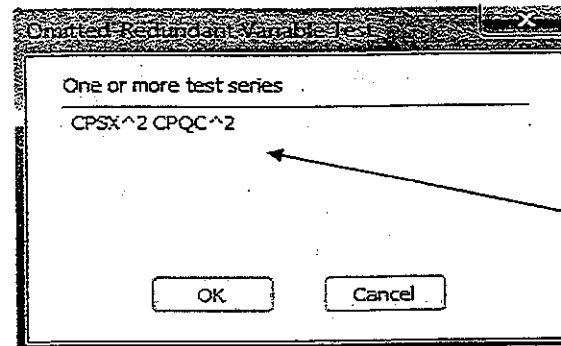
$H_0$  : nên bỏ biến CPQC ra khỏi mô hình

$H_1$  : không nên bỏ biến CPQC ra khỏi mô hình

Kết quả từ thống kê F cho thấy bác bỏ  $H_0$ . Vậy không nên loại biến CPQC.

*Ngoài ra việc kiểm định đưa thêm biến (mô hình ban đầu bỏ sót biến) vào trong mô hình cũng thực hiện tương tự với lệnh sau:*

[Equation] Views → Coefficient Tests → Omitted Variables–Likelihood Ratio ... Ví dụ ta đưa thêm 2 biến  $CPSX^2$  và  $CPQC^2$



Khai báo biến muốn đưa thêm vào là  $CPSX^2$  và  $CPQC^2$  rồi nhấn OK. Lưu ý là hai biến cách nhau một khoảng trắng

Đây là một nội dung quan trọng trong kiểm định sai dạng mô hình. Ý tưởng của kiểm định này là khi ta đưa thêm biến vào mô hình và muốn biết các biến này có đóng góp có ý nghĩa vào việc giải thích sự thay đổi của biến phụ thuộc hay không. Giả thiết  $H_0$  của kiểm định này là các biến mới đưa thêm vào mô hình đồng thời không có ý nghĩa. Điều kiện để áp dụng kiểm định này cũng giống như kiểm định thừa biến ở trên.

Kết quả kiểm định:

Omitted Variables: CPSX^2 CPQC^2

F-statistic	3.032223	Prob. F(2,15)	0.0784
Log likelihood ratio	6.790729	Prob. Chi-Square(2)	0.0335

Test Equation:

Dependent Variable: DT  
 Method: Least Squares  
 Date: 04/23/14 Time: 23:44  
 Sample: 1 20  
 Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CPSX	8.251830	2.321700	3.554219	0.0029
CPQC	0.464542	1.036787	0.448060	0.6605
C	-4.674312	7.807956	-0.598660	0.5583
CPSX^2	-0.319602	0.156238	-2.045606	0.0587
CPQC^2	0.133416	0.072982	1.828072	0.0875

R-squared	0.938787	Mean dependent var	46.05000
Adjusted R-squared	0.922463	S.D. dependent var	18.55710
S.E. of regression	5.167305	Akaike info criterion	6.334898
Sum squared resid	400.5156	Schwarz criterion	6.583831
Log likelihood	-58.34898	Hannan-Quinn criter.	6.383492
F-statistic	57.51118	Durbin-Watson stat	2.245427
Prob(F-statistic)	0.000000		

Kết quả trên cho biết điều gì?

7. Có người cho rằng mô hình hồi quy đưa ra không phù hợp. Hãy cho biết nhận xét này đúng hay sai?

Để biết nhận xét này đúng hay sai, ta đi kiểm định sự phù hợp của hàm hồi quy.

$H_0$  : mô hình hồi quy đưa ra không phù hợp.

$H_1$  : mô hình hồi quy đưa ra là phù hợp.

Tiêu chuẩn kiểm định:

$$F_{qs} = \frac{R^2 / (k-1)}{(1-R^2) / (n-k)} = 90,381$$

$$F_{\alpha}(k-1; n-k) = F_{0,05}(3-1; 20-3) = 3,592$$

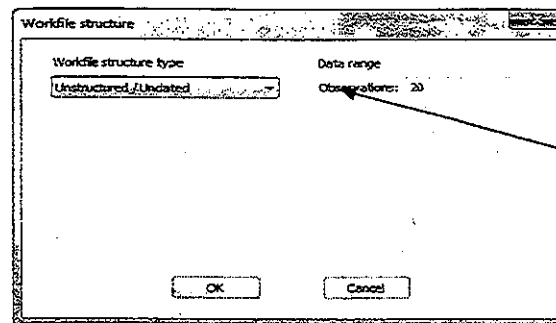
Ta có  $F_{qs} = 90,381 > 3,592$  nên đủ cơ sở bác bỏ  $H_0$ . Vậy hàm hồi quy phù hợp.

8. Sử dụng mô hình tuyến tính, tìm khoảng dự báo giá trị trung bình và giá trị riêng biệt cho DT khi  $CPSX = 10, CPQC = 2$ ?

Trước tiên, hồi quy mô hình tuyến tính của DT theo CPSX và CPQC.

Quan sát ứng với  $CPSX = 10, CPQC = 2$  chưa có trong mẫu, do đó ta cần mở rộng mẫu để nhập các giá trị này như sau:

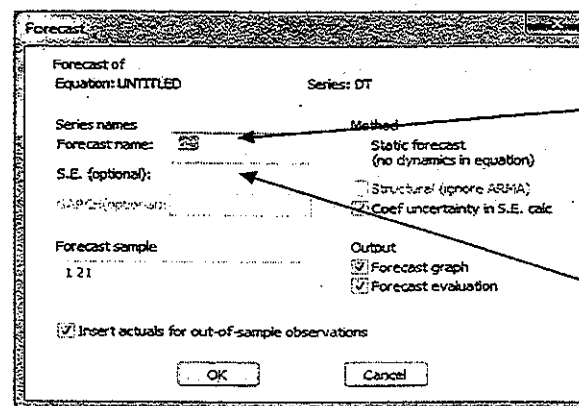
[Workfile]Proc → Structure/Resize Current Page... cửa sổ sau xuất hiện



Sửa lại số quan sát là 21 nếu muốn nhập thêm giá trị cho 1 quan sát. Sau đó chọn OK và xác nhận lại lần nữa chọn YES

Nếu trong mẫu đã có quan sát nào đó có giá trị  $CPSX = 10, CPQC = 2$  thì ta không cần thực hiện bước trên.

Sau đó, ta nhập giá trị cho các biến CPSX và CPQC. Tiếp theo, ta tìm dự báo điểm  $\widehat{DT}_0$  bằng cách từ cửa sổ Equation chọn Forecast sẽ xuất hiện cửa sổ sau:



Đặt tên biến cho dự báo điểm, ở đây là DTF, để lưu giá trị

Đây chính là  $se(Y_0 - \widehat{Y}_0)$ , để lưu giá trị này ta có thể đặt tên cho nó, chẳng hạn se\_rb

Thực hiện xong các thao tác trên trong cửa sổ Workfile sẽ xuất hiện 2 biến mới là DTF và SE\_RB. Mở biến DTF và nhìn vào quan sát thứ 21 ta thấy dự báo điểm của doanh thu khi chi phí sản xuất là 10 tỉ VNĐ và chi phí quảng cáo là 2 tỉ VNĐ là 46,96396 tỉ VNĐ.

Ta tạo biến se\_tb (để tính  $se(\widehat{Y}_0)$ ) từ cửa sổ lệnh như sau:

$$\text{genr se\_tb} = \text{sqr}(\text{se\_rb}^2 - 5.7519^2)$$

Tìm giá trị tới hạn Student mức  $\alpha/2 = 0,025$  và bậc tự do  $n - k = 18$ :

$$\text{scalar } t = @qtdist(0.975, 18)$$

Kết quả  $t = t_{0,025}(18) = 2,1009$ .

Tiếp theo, để tính các cận trên, cận dưới cho các khoảng dự báo ta có thể tạo ra các biến này từ cửa sổ lệnh lần lượt như sau:

genr cd\_tb = dtf - se\_tb\*t

genr ct\_tb = dtf + se\_tb\*t

genr cd\_rb = dtf - se\_rb\*t

genr ct\_rb = dtf + se\_rb\*t

Khi đó, trong cửa sổ Workfile sẽ có các biến cd\_tb, ct\_tb, cd\_rb, ct\_rb. Mở giá trị của nhóm biến này theo thứ tự như trên ta được bảng kết quả sau:

obs	CD_TB	CT_TB	CD_RB	CT_RB
1	18.71986	27.70991	10.32162	36.10805
2	30.95044	37.65475	21.76198	46.84321
3	35.52542	42.56477	26.45866	51.63153
4	14.94279	24.79924	6.820438	32.92159
5	70.46201	81.64376	62.73793	89.36784
6	55.26100	62.09535	46.12002	71.23633
7	19.73342	31.66452	12.22240	39.17554
8	51.34262	61.73003	43.38334	69.68951
9	39.02672	47.41218	30.42847	56.01043
10	30.45254	38.21096	21.64008	47.02342
11	39.88066	54.84863	33.15056	61.57873
12	72.67189	85.88829	65.50699	93.05320
13	22.49350	31.93503	14.24061	40.18792
14	48.48178	57.51189	40.12902	65.96465
15	50.66927	57.71192	41.60370	66.77749
16	60.92449	68.94746	52.20325	77.66869
17	21.89505	30.47919	13.36323	39.01101
18	41.44707	47.10443	31.86480	55.68670
19	41.33319	47.84476	32.07377	57.10418
20	64.10836	73.82075	55.94101	81.98910
21	40.80729	53.12062	33.40171	60.52621

Cận trên, cận dưới các khoảng dự báo cần tìm nằm ở dòng quan sát thứ 21

Vậy khoảng dự báo cho giá trị trung bình, giá trị riêng biệt của DT khi  $CPSX = 10, CPQC = 2$  là:

$$(40,80729 < E(DT|_{CPSX=10,CPQC=2}) < 53,12062)$$

$$(33,40171 < DT_{0|_{CPSX=10,CPQC=2}} < 60,52621)$$

Cũng với các bước thực hiện như vậy, ta có thể dự báo cùng lúc cho nhiều giá trị khác nhau.

### 3.3. Bài tập

**Bài 3.1.** Cho các số liệu sau: GDP – tổng sản phẩm quốc dân; K – vốn; đơn vị tính: triệu đôla Đài Loan (GDP, K tính theo giá cố định).

Dependent Variable: GDP  
Method: Least Squares  
Sample: 1 20  
Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
K	11342.06	2888.403	3.926759	0.0010
C	43746.70	19361.78	2.259436	0.0365
R-squared		Mean dependent var		109468.7
Adjusted R-squared	0.431469	S.D. dependent var		57734.42
S.E. of regression	43532.34	Akaike info criterion		24.29504
Sum squared resid	3.41E+10	Schwarz criterion		24.39461
Log likelihood	-240.9504	Hannan-Quinn criter.		24.31447
F-statistic	15.41944	Durbin-Watson stat		2.107725
Prob(F-statistic)	0.000989			

- Viết hàm hồi quy tổng thể và hàm hồi quy mẫu?
- Có thể nói rằng khi vốn thay đổi thì GDP thay đổi hay không?
- Tính  $R^2$  bằng các cách khác nhau?
- Thực hiện kiểm định sự phù hợp của mô hình bằng các cách khác nhau.

5. Khi K giảm 1 thì GDP giảm trong khoảng nào?
6. Khi K tăng 2 thì GDP tăng trong khoảng nào?
7. Khi vốn tăng thêm một triệu đôla thì GDP tăng tối thiểu bao nhiêu?
8. Có phải khi vốn tăng 1 triệu thì GDP tăng 10000 triệu hay không?
9. Vốn giảm 1 triệu thì GDP giảm nhiều hơn 1100 triệu không?

**Bài 3.2.** Cho các số liệu sau: CPI – chỉ số giá tiêu dùng (năm 1994 là 100), GDP – tổng sản phẩm quốc nội (tỷ USD). Số liệu được lấy từ quý I năm 2000 đến quý IV năm 2010 của Mỹ. (trang www.forecast.org). LOG là logarit cơ số tự nhiên ( $\ln$ ) của các biến tương ứng. Kết quả hồi quy như sau:

Dependent Variable: LOG(CPI)

Sample: 2000Q1 2010Q4

Included observations: 44

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(GDP)		0.070599	18.16263	0.0000
C	-6.805419	0.665080	-10.23248	0.0000

R-squared		Mean dependent var	5.273941
Adjusted R-squared	0.884372	S.D. dependent var	0.082302
S.E. of regression	0.027986	Akaike info criterion	-4.269822
Sum squared resid	0.032896	Schwarz criterion	-4.188723
Log likelihood	95.93609	Hannan-Quinn criter.	-4.239747
F-statistic	329.8812	Durbin-Watson stat	0.112344
Prob(F-statistic)	0.000000		

1. Viết hàm hồi quy tổng thể và hàm hồi quy mẫu?
2. Viết hàm số kinh tế ban đầu và cho biết ý nghĩa của các kết quả ước lượng nhận được?
3. Tính hệ số xác định bằng các cách khác nhau và cho biết ý nghĩa?

4. Khi GDP tăng 1% thì CPI tăng 1,3% đúng không? Tăng 1,2% đúng không? Tăng 1,25% đúng không? Tăng 1,32% đúng không?
5. Khi GDP giảm 1% thì CPI giảm tối đa bao nhiêu?

**Bài 3.3.** Cho các số liệu về Y – mức tiêu dùng thực phẩm (1 USD); X – thu nhập (1 USD) và N – số người trong hộ. Tiến hành khảo sát 38 hộ gia đình ở một địa phương thu được kết quả sau đây:

Dependent Variable: Y

Included observations: 38

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X	0.160789	0.039500	4.070588	0.0003
N	1527.369	462.5472	3.302083	0.0022
C	670.1653	2977.214	0.225098	0.8232

R-squared	0.433792	Mean dependent var	15532.24
Adjusted R-squared	0.401437	S.D. dependent var	6189.981
S.E. of regression	4788.992	Akaike info criterion	19.86168
Sum squared resid	8.03E+08	Schwarz criterion	19.99097
F-statistic	13.40738	Durbin-Watson stat	1.743033
Prob(F-statistic)	0.000048		

Cho biết  $Cov(X,N)$  là 0,459181

1. Viết các dạng mô hình hồi quy trên tổng thể và trên mẫu.
2. Các biến độc lập có ảnh hưởng tới mức tiêu dùng thực phẩm hay không?
3. Khi cả thu nhập và số người cùng tăng một đơn vị thì mức tiêu dùng thực phẩm thay đổi trong khoảng nào?
4. Khi hộ gia đình tăng thêm một người thì mức tiêu dùng tăng tối đa bao nhiêu?
5. Nếu tăng thêm 2 người thì mức tiêu dùng tăng trong khoảng nào?

- Có phải khi số người trong hộ gia đình tăng thêm 1 thì mức tiêu dùng thực phẩm tăng 2000 USD.
- Khi thực hiện hồi quy Y theo X ta thu được  $R^2 = 0,257$ . Có thể nói rằng việc đưa thêm biến N vào là phù hợp hay không?

**Bài 3.4.** Cho các số liệu sau: GDP; K – vốn; L – ngày lao động; đơn vị tính: triệu đôla Đài Loan (GDP, K tính theo giá cố định). Hồi quy thu được kết quả sau:

Dependent Variable: GDP  
Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
L	17662.45	4533.201	3.896242	0.0012
K	10751.92	2165.515	4.965061	0.0001
C	-21717.59	22180.83	-0.979116	0.3413
Adjusted R-squared	0.681997	S.D. dependent var	57734.42	
S.E. of regression	32557.46	F-statistic	21.37391	
Sum squared resid	1.80E+10			

- GDP là bao nhiêu khi vốn là 100 đơn vị và lao động là 50 đơn vị
- Tính hệ số xác định bội bằng các cách khác nhau?
- Phải chăng các biến độc lập không giải thích được cho sự biến động của GDP.
- Khi lao động không đổi, vốn tăng thêm một đơn vị thì GDP tăng tối đa bao nhiêu?
- Khi vốn không đổi, giảm đi một lao động thì GDP thay đổi khoảng bao nhiêu?

- Khi lao động không đổi, vốn tăng thêm một đơn vị thì GDP tăng 100 đơn vị phải không?
- Dùng kiểm định thu hẹp hồi quy để đánh giá xem việc đưa thêm lao động vào mô hình có phù hợp hay không biết rằng với mô hình GDP phụ thuộc vào K thì có  $R^2 = 0.461391$
- Khi vốn không đổi, giảm đi một ngày lao động thì GDP giảm ít hơn 100 đơn vị phải không?

**Bài 3.5.** Cho các số liệu đối với ngành dịch vụ của Mỹ từ năm 1948 – 1976 như sau:

Q: mức đầu; K – đầu vào vốn ; L – đầu vào lao động.

Year	Quantity	Labor	Capital
1948	51.4	75.8	49.7
1949	51.2	75.2	50.4
1950	52.7	74.3	51.3
1951	53.8	73.9	52.2
1952	55.2	74.9	52.3
1953	56.1	75.4	52.6
1954	56.1	75.5	53
1955	59.2	77	54.4
1956	62.4	78.6	56.7
1957	64.4	80.5	59.2
1958	66.2	80.7	61.4
1959	71.1	83.4	64.2
1960	71.1	85.8	67.4
1961	74.7	85.9	70.8
1962	77.4	89.3	75
1963	81.1	93	79.7
1964	86	94.5	84.2
1965	90	97	89.5
1966	94.6	99.3	95.4

1967	100	100	100
1968	103.2	101.2	107.2
1969	108.8	103.6	114.8
1970	113.4	106.5	119
1971	115.5	106.2	122.9
1972	124.9	110.5	127.4
1973	134.1	115.9	132.3
1974	136.9	119.5	136.5
1975	136.1	117	139.3
1976	143.7	120.5	142.3

Nguồn: Kendrick, John W. and Elliot S. Grossman (1980). Productivity in the United States: Trends and Cycles. Baltimore: Johns Hopkins Press. p. 140

Nhập và lưu số liệu thành file **ch3.5.wf1**

Kết quả ước lượng:

Mô hình (1)

(1) Dependent Variable: LOG(Q)  
Method: Least Squares  
Sample: 1948 1976  
Included observations: 29

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(K)	0.517556	0.126236	4.099917	0.0004
LOG(L)	0.892611	0.285167	3.130132	0.0043
C	-1.892169	0.739340	-2.559268	0.0167

R-squared	0.994487	Mean dependent var	4.396239
Adjusted R-squared	0.994063	S.D. dependent var	0.341824
S.E. of regression	0.026338	Akaike info criterion	-4.337908
Sum squared resid	0.018036	Schwarz criterion	-4.196464
Log likelihood	65.89967	Hannan-Quinn criter.	-4.293610
F-statistic	2345.124	Durbin-Watson stat	0.541266
Prob(F-statistic)	0.000000		

- Viết hàm số kinh tế ban đầu.
- Cho biết ý nghĩa của các ước lượng nhận được.
- Khi lao động tăng 1% thì Q thay đổi khoảng bao nhiêu?
- Vốn giảm 1% thì mức đầu ra giảm tối đa bao nhiêu %?
- Nguồn vốn tăng lên 1,7 lần so với trước thì mức đầu ra có tăng tương ứng là 1,7 lần hay không?
- K hay L tác động tới Q mạnh hơn.

Cho biết  $cov(\hat{\beta}_{\log(K)}, \hat{\beta}_{\log(L)}) = -0,035798$ .

Bài 3.6. Vẫn sử dụng tập số liệu **ch3.5**. Ước lượng bằng Eviews các mô hình sau:

Mô hình (2)

(2) Dependent Variable: LOG(Q)  
Method: Least Squares  
Sample: 1948 1976  
Included observations: 29

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(K)	0.517556	0.126236	4.099917	0.0004
LOG(L)	0.892611	0.285167	3.130132	0.0043
C	-1.892169	0.739340	-2.559268	0.0167

R-squared	0.994487	Mean dependent var	4.396239
Adjusted R-squared	0.994063	S.D. dependent var	0.341824
S.E. of regression	0.026338	Akaike info criterion	-4.337908
Sum squared resid	0.018036	Schwarz criterion	-4.196464
F-statistic	2345.124	Durbin-Watson stat	0.541266
Prob(F-statistic)	0.000000		

Mô hình (3)

(3) Dependent Variable: LOG(Q)  
 Method: Least Squares  
 Sample: 1948 1976  
 Included observations: 29

Variable	Coefficient	Std. Error	t-Statistic	Prob.
K	0.002328	0.001996	1.166333	0.2541
L	0.017077	0.004148	4.117156	0.0003
C	2.625906	0.215062	12.21000	0.0000
R-squared	0.985920	Mean dependent var		4.396239
Adjusted R-squared	0.984837	S.D. dependent var		0.341824
F-statistic	910.3271	Durbin-Watson stat		0.370832
Prob(F-statistic)	0.000000			

Mô hình (4)

(4) Dependent Variable: Q  
 Method: Least Squares  
 Sample: 1948 1976  
 Included observations: 29

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(K)	23.33863	20.89233	1.117091	0.2742
LOG(L)	126.4430	47.19585	2.679112	0.0126
C	-586.3430	122.3625	-4.791850	0.0001
R-squared	0.980297	Mean dependent var		85.90690
Adjusted R-squared	0.978781	S.D. dependent var		29.92455
S.E. of regression	4.359007	Akaike info criterion		5.880063
Sum squared resid	494.0246	Schwarz criterion		6.021507
Log likelihood	-82.26091	Hannan-Quinn criter.		5.924362
F-statistic	646.7935	Durbin-Watson stat		0.233668
Prob(F-statistic)	0.000000			

1. Viết các hàm kinh tế ban đầu của các mô hình trên.
2. Cho biết ý nghĩa của các mô hình.
3. Có thể so sánh  $R^2$  giữa các mô hình được không?
4. Cho biết mô hình nào là tốt nhất?

Bài 3.7. Cho các số liệu sau: (tập số liệu ch3.7)

- Y: MORTAL = infant mortality rate (per 1000 live births): tỷ lệ tử vong trẻ sơ sinh (trên 1000 trẻ đẻ sống) %
- $X_2$ : CALORIE = daily per capita calorie consumption: bình quân lượng calo mà một người tiêu thụ hàng ngày (kcal/người/ngày) (kcal: kilocalorie).
- $X_3$ : HS = percentage of age group enrolled in secondary education (female): tỷ lệ nhóm tuổi đi học ở bậc trung học (nữ) (%)
- $X_4$ : POPPHYS = population per physician: tỷ lệ bác sĩ trong dân
- $X_5$ : POPNURS = population per nursing person: tỷ lệ y tá điều dưỡng

Kết quả ước lượng được như sau:

Dependent Variable: Y  
 Method: Least Squares  
 Sample(adjusted): 1 99  
 Included observations: 99 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
$X_2$	-0.029568	0.009387	-3.150013	0.0022
$X_3$		0.219195	-5.616014	0.0000
$X_4$	0.000602	0.000195	3.088408	0.0026
$X_5$		0.000596	2.170630	0.0325
C	189.0972	20.15027	9.384350	0.0000
R-squared		Mean dependent var		104.4141
Adjusted R-squared	0.773957	S.D. dependent var		58.48653

S.E. of regression	27.80679	Akaike info criterion	9.537623
Sum squared resid	72682.47	Schwarz criterion	9.668689
Log likelihood	-467.1123	F-statistic	84.88668
Durbin-Watson stat	2.294874	Prob(F-statistic)	0.000000

Cho ma trận hiệp phương sai của các hệ số hồi quy ước lượng được từ mô hình trên như sau:

	X2	X3	X4	X5	C
X2	8.81E-05	-0.001477	1.92E-07	-6.93E-08	-0.180214
X3	-0.001477	0.048046	1.00E-05	2.30E-05	2.265392
X4	1.92E-07	1.00E-05	3.80E-08	-3.39E-08	-0.001053
X5	-6.93E-08	2.30E-05	-3.39E-08	3.55E-07	-0.001149
C	-0.180214	2.265392	-0.001053	-0.001149	406.0333

- Hãy điền vào các chỗ còn trống trong bảng kết quả hồi quy?
- Cho biết khi X3 và X5 cùng tăng 1 đơn vị thì Y thay đổi tối đa là bao nhiêu?
- Giữa X2 và X4, biến nào tác động tới Y mạnh hơn?
- Theo kết quả ước lượng nhận được dưới đây thì có nên bỏ X4 và X5 ra khỏi mô hình hay không?

Dependent Variable: Y  
Method: Least Squares  
Sample (adjusted): 1 99  
Included observations: 99 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X2	-0.033185	0.010164	-3.265077	0.0015
X3	-1.569430	0.223896	-7.009631	0.0000
C	217.5342	20.76096	10.47804	0.0000

R-squared	0.737409	Mean dependent var	104.4141
Adjusted R-squared	0.731938	S.D. dependent var	58.48653
S.E. of regression	30.28122	Akaike info criterion	9.688767
Sum squared resid	88027.44	Schwarz criterion	9.767407
Log likelihood	-476.5940	Hannan-Quinn criter.	9.720585
F-statistic	134.7936	Durbin-Watson stat	2.152178
Prob(F-statistic)	0.000000		

Bài 3.8. Cho các số liệu sau: (tập số liệu ch3.8)

P = price of lemonade, giá nước chanh

Qd = lượng cầu nước chanh

temp = nhiệt độ trung bình trong ngày

Psoda = giá nước soda

Bảng số liệu sau đây:

P	Qd	Temp	Psoda
137	0.2	82	0.75
117	1	88	0.75
106	0.7	70	0.6
108	0.85	70	0.5
115	0.25	62	1.2
111	1	57	1
101	0.25	51	0.3
125	0.45	88	0.55
109	0.15	58	0.35
99	0.9	67	0.65
115	0.7	94	0.7
124	0.2	74	1.2
123	0.55	85	0.95
107	0.1	51	0.4
121	0.1	94	0.25
123	0.3	89	0.9
128	0.75	86	0.85
120	0.75	87	0.5
112	0.55	53	0.25
110	0.2	57	0.7



110	1	60	0.75
116	0.8	80	0.9
108	0.85	64	0.6
113	0.85	51	1.2
106	0.55	60	0.55
106	0.4	56	0.85
128	0.15	87	0.4
114	0.9	64	1.1
122	0.3	53	0.85
100	0.75	57	0.9

1. Hãy tìm dạng mô hình hồi quy tốt nhất cho mỗi quan hệ giữa lượng cầu nước chanh với các yếu tố còn lại?
2. Với dạng mô hình tốt nhất đó, hãy cho biết khi giá nước chanh tăng 1 thì lượng cầu nước chanh giảm bao nhiêu? Tối đa bao nhiêu? Tối thiểu bao nhiêu?
3. Có phải yếu tố nhiệt độ tác động tới lượng cầu nước chanh mạnh hơn so với giá nước soda?
4. Yếu tố nhiệt độ tác động thế nào tới lượng cầu nước chanh? Nếu trời nóng thêm 1 độ thì lượng cầu nước chanh tăng tối thiểu là bao nhiêu?
5. Dự báo mức cầu nước chanh khi nhiệt độ là 100 đơn vị, giá bán nước chanh là 150 đơn vị và giá soda là 1,5 đơn vị.
6. Dùng kiểm định cần thiết để xem xét mô hình có thiếu biến hay thừa biến không?

## Chương 4:

### PHÂN TÍCH HỒI QUY VỚI BIẾN ĐỊNH TÍNH

#### 4.1. Mô hình hồi quy với biến giả

Trong những mô hình hồi quy mà chúng ta đã xét ở trên, các biến đều là biến định lượng như thu nhập, sản lượng, giá cả hàng hóa,... Thực tế cho thấy nhiều trường hợp không thể giải thích một cách thuận tụy bằng biến định lượng mà phải giải thích bằng các biến định tính như giới tính, hoàn cảnh gia đình, trình độ học vấn, quốc tịch,...

Để đưa biến định tính vào mô hình, người ta lượng hóa biến định tính bằng biến giả nhận hai giá trị 0 và 1.

Chẳng hạn, biến định tính: khu vực kinh tế (tư nhân, nhà nước).

Để chuyển biến Khu vực kinh tế thành biến định lượng ta đặt:

$$\begin{aligned} \text{TN} &= 1 \text{ nếu quan sát thuộc khu vực kinh tế tư nhân} \\ &= 0 \text{ nếu quan sát thuộc khu vực kinh tế khác} \end{aligned}$$

Biến định tính: Vùng miền: Bắc ; Trung ; Nam

$$\begin{aligned} \text{Bắc} &= 1 \text{ nếu quan sát ở miền Bắc} \\ &= 0 \text{ nếu quan sát ở miền khác} \end{aligned}$$

$$\begin{aligned} \text{Trung} &= 1 \text{ nếu quan sát ở miền Trung} \\ &= 0 \text{ nếu quan sát ở miền khác} \end{aligned}$$

**Một số lưu ý khi đặt biến giả:**

- Biến giả chỉ nhận hai giá trị 0 và 1 (hai số không có thứ tự và không có đơn vị đo).
- Để phân biệt 2 phạm trù nam hay nữ ta dùng một biến giả, khi có m phạm trù ta dùng m - 1 biến giả. Số biến

giả thấp hơn số phạm trù là 1 để tránh hiện tượng đa cộng tuyến hoàn hảo.

- Đối với biến định tính có nhiều hơn 2 phạm trù, ta gọi phạm trù tương ứng được gán giá trị 0 là phạm trù cơ sở.
- Phạm trù nào cũng phải có giá trị riêng biệt.
- Trong lý thuyết biến giả thường được đặt tên là D - dummy. Tuy nhiên, trong thực hành, để thuận lợi khi phân tích bằng Eviews, biến giả thường được đặt tên trùng với thuộc tính nhận giá trị 1.
- Các hệ số góc gắn với biến giả được gọi là các hệ số chênh lệch, ý nghĩa của nó đo lường mức chênh lệch giữa các phạm trù chứ không phải đo độ lớn của mối quan hệ giữa biến giả và biến phụ thuộc.

#### 4.2. Mô hình với biến độc lập là định lượng và định tính

Xét mô hình với biến định lượng Y phụ thuộc biến độc lập định lượng X có dạng:

$$Y_i = \beta_1 + \beta_2 X_i + u_i$$

Giả sử có biến định tính A có hai phạm trù  $A_1$  và  $A_2$  (hay  $\bar{A}_2$ ) tác động đến mối quan hệ X - Y thông qua việc tác động đến hệ số chặn và hệ số góc.

Để phân tích, ta đặt biến giả D như sau:

$$D_i = 1 \text{ nếu quan sát } i \text{ thuộc phạm trù } A_1$$

$$D_i = 0 \text{ nếu quan sát } i \text{ thuộc phạm trù ngược lại}$$

Các trường hợp tác động có thể xảy ra:

+ *Biến định tính tác động đến hệ số chặn:*

$$\text{Mô hình có dạng: } Y_i = \beta_1 + \beta_2 X_i + \beta_3 D_i + u_i$$

$$\text{Tại } A_1: D = 1: Y_i = (\beta_1 + \beta_3) + \beta_2 X_i + u_i$$

$$\text{Tại } A_2: D = 0: Y_i = \beta_1 + \beta_2 X_i + u_i$$

+ *Biến định tính tác động đến hệ số góc:*

$$\text{Mô hình có dạng: } Y_i = \beta_1 + \beta_2 X_i + \beta_3 D_i X_i + u_i$$

$$\text{Tại } A_1: D = 1: Y_i = \beta_1 + (\beta_2 + \beta_3) X_i + u_i$$

$$\text{Tại } A_2: D = 0: Y_i = \beta_1 + \beta_2 X_i + u_i$$

+ *Biến định tính tác động đến cả hai hệ số:*

$$\text{Mô hình có dạng: } Y_i = \beta_1 + \beta_2 X_i + \beta_3 D_i + \beta_4 D_i X_i + u_i$$

$$\text{Tại } A_1: D = 1: Y_i = (\beta_1 + \beta_3) + (\beta_2 + \beta_4) X_i + u_i$$

$$\text{Tại } A_2: D = 0: Y_i = \beta_1 + \beta_2 X_i + u_i$$

Tùy sự tác động của biến định tính đến hệ số chặn hoặc hệ số góc mà xác định mô hình phù hợp.

#### 4.3. Một số ứng dụng của biến giả

##### 4.3.1. Kiểm định sự thay đổi cấu trúc của mô hình

Khi hồ qui mẫu số liệu theo cuối thời gian, do trải qua các thời kỳ khác nhau nên có khả năng xảy ra những thay đổi về cấu trúc trong mối quan hệ giữa biến phụ thuộc Y và các biến độc lập X. Qua đó, giá trị của các tham số hồi quy có thể bị biến động. Hay khi tổng thể gồm hai nhóm mang đặc tính khác nhau và ta muốn biết các hệ số hồi quy của hai nhóm này có như nhau (tức là các hệ số là ổn định) hay không.

##### *Kiểm định Chow*

Giả sử ta khảo sát một mẫu số liệu có n quan sát (đã được sắp xếp theo thời gian hay sắp xếp thành hai nhóm theo các đặc tính). Mẫu số liệu này có thể được chia thành hai mẫu nhỏ với số quan sát là  $n_1$  (ứng với nhóm thứ nhất) và  $n - n_1$  (ứng với nhóm thứ hai).

Xét các mô hình:

$$Y_i = \alpha_1 + \alpha_2 X_i + u_i \quad (i = 1, 2, \dots, n_1) \quad (4.1)$$

$$Y_i = \beta_1 + \beta_2 X_i + u_i \quad (i = n_1 + 1, n_1 + 2, \dots, n) \quad (4.2)$$

$$Y_i = \gamma_1 + \gamma_2 X_i + u_i \quad (i = 1, 2, \dots, n) \quad (4.3)$$

Giả thiết kiểm định:

$$H_0: \alpha_1 = \beta_1, \alpha_2 = \beta_2$$

$$H_1: \text{ít nhất một cặp hệ số } \alpha_j \neq \beta_j, j = 1, 2$$

Hay  $H_0$ : cấu trúc của hai mô hình (4.1) và (4.2) như nhau

$H_1$ : cấu trúc của hai mô hình (4.1) và (4.2) khác nhau

Lần lượt hồi quy các mô hình 4.1, 4.2 và 4.3 theo các mẫu tương ứng thu được  $RSS_1, RSS_2$  và  $RSS$ .

Tính tiêu chuẩn kiểm định:

$$F_{qs} = \frac{[RSS - (RSS_1 + RSS_2)]/k}{(RSS_1 + RSS_2)/(n - 2k)}$$

Nếu  $F_{qs} > F_{0,05}(k, n - 2k)$  thì bác bỏ giả thiết  $H_0$ .

**Kiểm định bằng phương pháp dùng biến giả**

Đặt biến giả

$$D = 1 \text{ nếu thuộc nhóm thứ nhất}$$

$$D = 0 \text{ nếu thuộc nhóm thứ hai}$$

Xét mô hình cho toàn bộ mẫu:

$$Y_i = \beta_1 + \beta_2 X_i + \beta_3 D_i + \beta_4 D_i X_i + u_i$$

Thực hiện kiểm định các giả thiết  $\beta_3 = 0, \beta_4 = 0$  để đưa ra kết luận.

### 4.3.2. Biến giả trong phân tích mùa

Nhiều chuỗi thời gian trong kinh tế có tính chất thời vụ rất rõ như doanh số bán của các mặt hàng điện lạnh thường tăng cao

vào dịp hè, doanh số bán ra của các cửa hàng văn phòng phẩm thường tăng cao vào những ngày đầu năm học,...

Thông thường ta muốn loại nhân tố mùa khỏi chuỗi thời gian để có thể tập trung vào các thành phần khác của chuỗi thời gian. Phương pháp sử dụng biến giả có thể sử dụng được để thực hiện mục đích trên.

Chẳng hạn, xét lợi nhuận  $Y$  theo doanh thu  $X$  của một số công ty trong một số năm. Giả sử chuỗi số liệu có tính chất mùa vụ theo từng quý.

Ta cần đánh giá ảnh hưởng của  $X$  lên  $Y$  theo tính xu hướng, tức là loại bỏ yếu tố mùa vụ. Ta đặt các biến giả như sau:

$D_1 = 1$  nếu quan sát ở quý I,  $D_1 = 0$  nếu quan sát ở các quý khác

$D_2 = 1$  nếu quan sát ở quý II,  $D_2 = 0$  nếu quan sát ở các quý khác

$D_3 = 1$  nếu quan sát ở quý III,  $D_3 = 0$  nếu quan sát ở các quý khác

Quý IV là quý cơ sở.

Nếu chỉ xét đến sự khác biệt trong tung độ gốc do ảnh hưởng của yếu tố mùa thì ta dùng mô hình sau:

$$Y_i = \beta_1 + \beta_2 X_i + \beta_3 D_1 + \beta_4 D_2 + \beta_5 D_3 + u_i$$

Nếu xét cả sự khác biệt trong hệ số góc do ảnh hưởng của yếu tố mùa thì ta dùng mô hình có các biến tương tác sau đây:

$$Y_i = \beta_1 + \beta_2 X_i + \beta_3 D_1 + \beta_4 D_2 + \beta_5 D_3 + \beta_6 D_1 X_i + \beta_7 D_2 X_i + \beta_8 D_3 X_i + u_i$$

## 4.4. Hướng dẫn Eviews

### 4.4.1. Kiểm định Chow

Mục đích của kiểm định Chow là xem liệu có sự thay đổi về mặt cấu trúc của mô hình hồi quy (đối với hồi quy chuỗi thời gian) giữa các giai đoạn khác nhau (do thay đổi chính sách hoặc cú sốc kinh tế) hay không. Ta xét ví dụ trong tập số liệu CH4VD1.wf1 với các biến số sau đây:

PCE – PCE services (chỉ tiêu tiêu dùng cá nhân cho dịch vụ) (tỷ USD)

GDP – Gross domestic product (tổng sản phẩm quốc dân) (tỷ USD)

FFR - Federal Funds Rate (lãi suất quỹ dự trữ của Úc) (%)

M2 – lượng cung tiền (tỷ USD)

Tiến hành khảo sát số liệu kinh tế Úc từ quý I năm 1983 đến quý I năm 2008.

Mô hình hồi quy cho toàn bộ thời kỳ:

$$PCE_t = \beta_1 + \beta_2 GDP_t + \beta_3 FFR_t + \beta_4 M2_t + u_t$$

Minh họa một phần bộ số liệu như sau:

C2-T	PCE - Services	Gross domestic product	Federal Funds Rate	Money M2
1983-I	2145	5252	9.29	1898
1983-II	2176	5370	8.65	1990
1983-III	2208	5476	8.80	2042
1983-IV	2233	5590	9.46	2077
1984-I	2241	5700	9.43	2121
1984-II	2263	5796	9.69	2157
1984-III	2292	5851	10.56	2204
1984-IV	2323	5902	11.39	2236
1985-I	2368	5955	9.27	2291
1985-II	2389	6008	8.48	2351
...	...	...	...	...

Kết quả ước lượng mô hình:

Dependent Variable: PCE

Method: Least Squares

Date: 05/21/13 Time: 07:49

Sample: 1983Q1 2008Q1

Included observations: 101

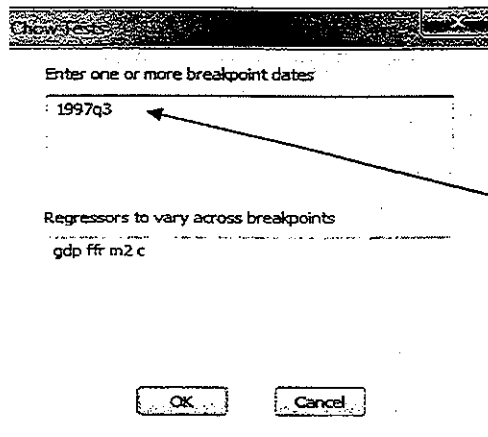
Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP	0.370003	0.007799	47.44488	0.0000
FFR	-11.13448	1.651324	-6.742761	0.0000
M2	0.033806	0.009395	3.598469	0.0005
C	194.9818	35.11981	5.551904	0.0000

R-squared	0.998576	Mean dependent var	3376.426
Adjusted R-squared	0.998532	S.D. dependent var	754.2616
S.E. of regression	28.89912	Akaike info criterion	9.604298
Sum squared resid	81010.45	Schwarz criterion	9.707867
Log likelihood	-481.0170	Hannan-Quinn criter.	9.646226
F-statistic	22674.33	Durbin-Watson stat	0.359162
Prob(F-statistic)	0.000000		

Trong thời gian từ quý I năm 1983 đến quý I năm 2008 có một mốc thời gian xảy ra cuộc khủng hoảng tài chính châu Á (còn gọi là khủng hoảng tiền tệ châu Á) xảy ra vào tháng 7 năm 1997 tại Thái Lan. Câu hỏi đặt ra là khủng hoảng này có ảnh hưởng đến cấu trúc của hàm hồi quy hay không, thời điểm này có phá vỡ cấu trúc về chỉ tiêu tiêu dùng cá nhân cho dịch vụ hay không. Để trả lời cho câu hỏi này có hai cách, một là thực hiện biến giả, hai là kiểm định Chow. Eviews thực hiện kiểm định Chow như sau:

[Equation] View → Stability Tests → Chow Breakpoint Test



Ta nghỉ ngơi thời điểm Quý 3 năm 1997 khi xảy ra khủng hoảng tài chính châu Á (xuất phát từ Thái Lan) thì nền kinh tế Úc cũng chịu ảnh hưởng. Để xem có sự thay đổi về cấu trúc của hàm hồi quy tại thời điểm này hay không ta gõ 1997Q3 vào Chow Test rồi nhấn OK

**Kết quả kiểm định Chow:**

**Chow Breakpoint Test: 1997Q3**  
**Null Hypothesis: No breaks at specified breakpoints**  
**Varying regressors: All equation variables**  
**Equation Sample: 1983Q1 2008Q1**

F-statistic	7.214117	Prob. F(4,93)	0.0000
Log likelihood ratio	27.29468	Prob. Chi-Square(4)	0.0000
Wald Statistic	28.85647	Prob. Chi-Square(4)	0.0000

Cặp giả thiết như sau:

- $H_0$ : hồi quy của hai thời kỳ trước và sau quý 3 năm 1997 là như nhau
- $H_1$ : hồi quy của hai thời kỳ trước và sau quý 3 năm 1997 là khác nhau

Kết quả kiểm định cho thấy giá trị  $Prob.F(4,93) = 0,0000 < 0,05$  (mức ý nghĩa 5%) cho nên ta bác bỏ  $H_0$  và thừa nhận  $H_1$ , có sự thay đổi về cấu trúc trong hàm hồi quy. Nói cách khác cuộc khủng hoảng tài chính Châu Á năm 1997 có ảnh hưởng đến cấu trúc của hàm hồi quy.

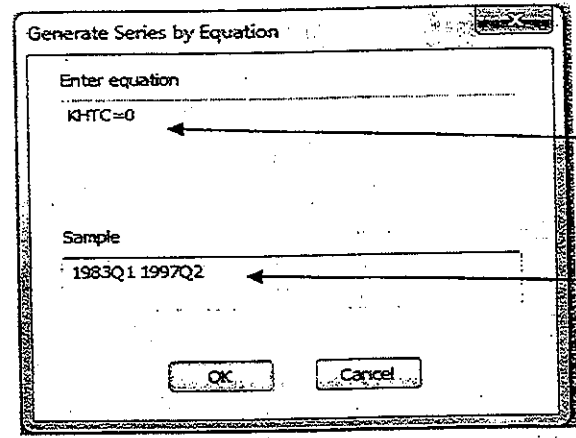
**4.4.2. Thực hiện kiểm định thay đổi về cấu trúc với phương pháp biến giả**

Ở trên ta đã dùng kiểm định Chow để xem xét sự thay đổi về cấu trúc, trong mục này ta sẽ xem xét cách đặt biệt giả:

$$KHTC = 1 \text{ nếu là thời điểm từ quý 3 năm 1997}$$

$$= 0 \text{ nếu là thời điểm trước quý 3 năm 1997}$$

Để tạo biến giả KHTC với các giá trị như trên, trong Eviews ta thực hiện như sau: từ cửa sổ **Workfile** chọn **Genr**



Đặt tên biến và gán giá trị cho biến KHTC

Giai đoạn mẫu ứng với giá trị gán ở trên

Tiếp tục gán giá trị 1 cho biến KHTC ở giai đoạn sau tương tự như trên.

Minh họa một phần bộ số liệu như sau:

C2-T	PCE - Services	Gross domestic product	Federal Funds Rate	Money M2	KHTC
...	...	...	...	...	...
1996-II	3345	8307	5.36	3675	0
1996-III	3366	8363	5.24	3721	0
1996-IV	3386	8470	5.31	3752	0
1997-I	3418	8544	5.28	3802	0
1997-II	3447	8658	5.28	3848	0

1997-III	3485	8773	5.52	3890	1
1997-IV	3521	8837	5.53	3943	1
1998-I	3562	8936	5.51	4015	1
1998-II	3602	8992	5.52	4092	1
1998-III	3642	9097	5.50	4164	1
...	...	...	...	...	...

Kết quả ước lượng mô hình có thêm biến giả như sau:

Dependent Variable: PCE  
Method: Least Squares  
Date: 04/24/14 Time: 10:32  
Sample: 1983Q1 2008Q1  
Included observations: 101

Variable	Coefficient	Std. Error	t-Statistic	Prob.
KHTC	558.9809	234.1512	2.387265	0.0190
GDP	0.334546	0.018440	18.14248	0.0000
FFR	-10.40685	2.375442	-4.381018	0.0000
M2	0.112913	0.031872	3.542714	0.0006
KHTC*GDP	-0.078048	0.045583	-1.712226	0.0902
KHTC*FFR	11.04797	4.092886	2.699310	0.0083
KHTC*M2	0.019312	0.047100	0.410022	0.6827
C	201.8420	53.75814	3.754631	0.0003

R-squared	0.998913	Mean dependent var	3376.426
Adjusted R-squared	0.998831	S.D. dependent var	754.2616
S.E. of regression	25.78376	Akaike info criterion	9.413261
Sum squared resid	61826.61	Schwarz criterion	9.620400
Log likelihood	-467.3697	Hannan-Quinn criter.	9.497117
F-statistic	12211.84	Durbin-Watson stat	0.370410
Prob(F-statistic)	0.000000		

Với phương pháp biến giả ta có thể thấy được sự thay đổi về cấu trúc của PCE là do biến độc lập FFR gây ra, còn các yếu tố như GDP và M2 không có sự thay đổi nhiều giữa hai thời kỳ (bạn đọc có thể dễ dàng kiểm tra). Phương pháp biến giả để kiểm tra sự thay đổi cấu trúc cũng đồng thời xác định được hai hàm hồi quy cho hai thời kỳ trước và sau quý 3 năm 1997.

#### 4.4.3. Ví dụ khác

**Nội dung:** xem xét sự phụ thuộc của giá bán nhà vào các yếu tố: diện tích, giá vàng tại thời điểm, hướng của ngôi nhà (đông nam, tây nam, đông bắc, tây bắc). Khảo sát 100 mảnh quảng cáo rao bán nhà tại Bình Dương.

Biến phụ thuộc: GN: Giá bán nhà (Đơn vị: triệu đồng)

Biến độc lập: DT: diện tích đất (m<sup>2</sup>)

GV: Giá vàng (triệu đồng/lượng)

Biến định tính: hướng của ngôi nhà (huong): đông nam (DN), tây nam (TN), đông bắc (DB), tây bắc (TB)

HDN = 1 nếu nhà ở hướng Đông Nam.

= 0 nếu nhà không ở hướng Đông Nam.

HTN = 1 nếu nhà ở hướng Tây Nam.

= 0 nếu nhà không ở hướng Tây Nam.

HDB = 1 nếu nhà ở hướng Đông Bắc.

= 0 nếu nhà không ở hướng Đông Bắc.

Bộ số liệu được lưu trong file CH4VD2.wf1.

Để tạo các biến giả như trên trong Eviews ta làm như sau:

+ Tạo biến HDN: từ cửa sổ lệnh ta lần lượt thực hiện các lệnh

Gener HDN = 0

Smpl if huong = "DN"

Gener HDN = 1

Smpl @all

+ Tạo các biến HTN, HDB được thực hiện tương tự.

Việc đặt biến giả cũng có thể làm trực tiếp trong Excel rồi copy số liệu qua Eviews. Minh họa một phần bộ số liệu:

Giá bán (triệu đồng)	Diện tích (m <sup>2</sup> )	Giá vàng (triệu đồng / lượng)	HDN	HTN	HDB	Hướng
140	36	37.12	0	1	0	Tây Nam
275	35	37.62	1	0	0	Đông Nam
275	65	37.57	0	1	0	Tây Nam
320	48	43.08	0	1	0	Tây Nam
325	48	37.91	0	1	0	Tây Nam
330	36	39.91	0	0	0	Tây Bắc
340	80	37.62	1	0	0	Đông Nam
350	36	46.00	1	0	0	Đông Nam
350	50	37.12	0	1	0	Tây Nam
360	50	37.12	0	1	0	Tây Nam
360	37	46.30	0	1	0	Tây Nam
375	70	45.95	0	1	0	Tây Nam
380	36	45.95	0	0	0	Tây Bắc
...	...	...	...	...	...	...

Nguồn: Quảng cáo trên Báo Thanh niên.

Việc ước lượng được thực hiện như thông thường. Kết quả ước lượng như sau:

Dependent Variable: GN  
Method: Least Squares  
Date: 04/24/14 Time: 09:04  
Sample: 1 100  
Included observations: 100

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DT	1.668641	0.308463	5.409540	0.0000
GV	16.34762	5.793881	2.821532	0.0058
HDN	-121.6253	57.55568	-2.113177	0.0372
HTN	-81.32406	48.46352	-1.678047	0.0967

HDB	-137.3393	73.12015	-1.878269	0.0634
C	-164.3959	258.3348	-0.636368	0.5261
R-squared	0.312336	Mean dependent var		639.7300
Adjusted R-squared	0.275758	S.D. dependent var		207.2318
S.E. of regression	176.3593	Akaike info criterion		13.24105
Sum squared resid	2923645.	Schwarz criterion		13.39736
Log likelihood	-656.0524	Hannan-Quinn criter.		13.30431
F-statistic	8.538918	Durbin-Watson stat		0.504885
Prob(F-statistic)	0.000001			

- Viết hàm hồi quy tổng thể và hàm hồi quy mẫu
- Có sự khác nhau về giá bán nhà giữa các hướng với hướng Tây Bắc hay không?

#### 4.5. Bài tập

Bài 4.1. Giá bán của một căn nhà phụ thuộc vào nhiều yếu tố, trong đó có:

- ✓ diện tích đất
- ✓ diện tích nhà
- ✓ vị trí mảnh đất (mặt tiền, hẻm chính, hẻm phụ)
- ✓ tuổi thọ căn nhà tính đến thời điểm bán
- ✓ tình trạng căn nhà tính đến thời điểm bán (tốt, còn sử dụng được, kém)
- ✓ hình dạng của mảnh đất (hình chữ nhật, hình vuông, nở hậu)

- Lập mô hình để phân tích các mối quan hệ trên?
- Nếu giá bán nhà nở hậu thường cao hơn nhà có hình vuông thì cần làm gì để kiểm tra nhận định này?

**Bài 4.2.** Hãy lập mô hình hồi quy diễn tả điều sau:

Lượng cam bán được phụ thuộc: giá cam, giá quýt, nơi bán (chợ sân bay, chợ khu 6, chợ đêm). Có người cho rằng ảnh hưởng ngược chiều của giá cam lên lượng cam bán được là tăng dần theo thứ tự. Hãy xây dựng mô hình và nêu cách kiểm tra nhận định trên.

**Bài 4.3.** Cho biết ý nghĩa của các hệ số hồi quy trong các mô hình sau đây:

1. Mô hình:  $\ln Y_i = \beta_1 + \beta_2 \ln X_i + \beta_3 D_i + \beta_4 D \ln X_i + u_i$

2. Mô hình:  $\ln Y_i = \beta_1 + \beta_2 \ln X_i + \beta_3 D_i + u_i$

**Bài 4.4.** Cho mô hình hồi quy sau đây:

$$CT = \beta_1 + \beta_2 TN + \beta_3 Nu + u$$

Trong đó CT và TN là chi tiêu và thu nhập của người lao động theo tháng, Nu là biến giả nhận giá trị 1 với lao động nữ và 0 với lao động nam. Nếu có người cho rằng cùng mức thu nhập, lao động nữ thường có mức chi tiêu tiết kiệm hơn so với nam giới thì cần làm gì để kiểm tra nhận định này?

**Bài 4.5.** Cho bộ số liệu sau, với các biến Y là mức tiêu dùng bia/người/năm; M là mức thu nhập/người/năm; Nu là giới tính (1 là nữ, 0 là nam). Hồi quy được kết quả sau đây:

Dependent Variable: Y  
Method: Least Squares  
Sample: 1 40  
Included observations: 40

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	213.9677	36.44385	5.871160	0.0000
M	0.001299	0.000509	2.549002	0.0151
Nu	-156.3140	39.59615	-3.947708	0.0003
R-squared	0.391681	Mean dependent var		191.5500
F-statistic	11.91167	Durbin-Watson stat		2.261267
Prob(F-statistic)	0.000101			

- Viết hàm hồi quy tổng thể và hàm hồi quy mẫu cho nam và nữ?
- Cho biết ý nghĩa của các hệ số góc?
- Hãy ước lượng mức tiêu dùng bia khi thu nhập tăng một đơn vị?
- Ta đưa thêm biến tương tác Nu\*M thì được mô hình mới

Dependent Variable: Y  
Method: Least Squares  
Sample: 1 40  
Included observations: 40

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	207.2111	50.12123	4.134198	0.0002
M	0.001445	0.000900	1.606762	0.1168
Nu	-146.6845	62.79076	-2.336084	0.0252
Nu*M	-0.000219	0.001098	-0.199363	0.8431
R-squared	0.392352	Mean dependent var		191.5500
F-statistic	7.748269	Durbin-Watson stat		2.268845
Prob(F-statistic)	0.000406			

- Cho biết khi mức thu nhập là 300 thì mức tiêu dùng bia của nam là bao nhiêu, nữ là bao nhiêu và giữa nam và nữ khác nhau bao nhiêu.
- Khi thu nhập cùng tăng một đơn vị thì mức tiêu dùng bia chênh lệch trong khoảng nào?

**Bài 4.6.** Cho các số liệu về CEO trong các doanh nghiệp trên tạp chí Forbes sau: (tập số liệu ch4.6)

Salary: lương và thưởng năm 1999

Tenure: nhiệm kỳ làm CEO, bằng 0 nếu ít hơn 6 tháng và bằng 1 nếu lâu hơn



Age: tuổi của CEO

Sales: doanh số bán hàng của hãng năm 1998

Profits: lợi nhuận của hãng năm 1998

Assets: tổng tài sản của hãng năm 1998

Kết quả ước lượng như sau:

Dependent Variable: SALARY

Method: Least Squares

Sample: 1 447

Included observations: 447

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AGE	5.492393	11.46086	0.479230	0.6320
ASSETS	0.007630	0.001326	5.754849	0.0000
SALES	0.014287	0.006614	2.160040	0.0313
TENURE	31.67279	9.465097	3.346272	0.0009
PROFITS	0.141302	0.068845	2.052471	0.0407
C	998.7095	623.6954	1.601277	0.1100

R-squared	0.243829	Mean dependent var	2027.517
Adjusted R-squared	0.240312	S.D. dependent var	1722.566
S.E. of regression	1501.390	Akaike info criterion	17.47950
Sum squared resid	9.94E+08	Schwarz criterion	17.53457
Log likelihood	-3900.669	Hannan-Quinn criter.	17.50121
F-statistic	29.21660	Durbin-Watson stat	2.014806
Prob(F-statistic)	0.000000		

1. Hãy cho biết biến Tensure có hưởng tới mức lương của CEO hay không?
2. Cho biết mức tăng tối đa của lương khi nhiệm kỳ làm CEO lớn hơn 6 tháng?

## Chương 5:

# KIỂM ĐỊNH VÀ LỰA CHỌN MÔ HÌNH

### 5.1. Một số khuyết tật của mô hình hồi quy

#### 5.1.1. Phương sai sai số (PSSS) thay đổi

Mô hình hồi quy:  $Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i$  (5.1)

**Giả thiết:** Phương sai các sai số ngẫu nhiên là đồng nhất (bằng nhau, thuần nhất hoặc đồng đều). Tức là  $Var(u_i) = \sigma^2$

Nếu  $Var(u_i) = \sigma_i^2 \rightarrow$  PSSS thay đổi (heteroscedasticity)

Cặp giả thiết của kiểm định khuyết tật:

$H_0$ : mô hình ban đầu có phương sai sai số bằng nhau (đồng đều)

$H_1$ : mô hình ban đầu có phương sai sai số thay đổi

#### 5.1.1.1. Kiểm định Glejer

Tư tưởng của dạng kiểm định này là dựa trên một cấu trúc giả định cụ thể nào đó của dạng thay đổi của phương sai. Người ta xem xét đồ thị phần dư và mô hình hoá sự thay đổi của nó, coi phương sai sai số có dạng hàm của một biến độc lập nào đó rồi đưa ra mô hình hồi quy phụ.

Sau khi thu được phần dư từ mô hình hồi quy được ước lượng bằng OLS, Glejer đề nghị hồi quy  $|e_i|$  theo biến X nào đó có liên quan tới  $\sigma_i^2$ . Glejer sử dụng các dạng hàm hồi quy phụ như sau:

STT	Mô hình hồi quy phụ	STT	Mô hình hồi quy phụ
1	$ e_i  = \alpha_1 + \alpha_2 X_i + v_i$	4	$ e_i  = \alpha_1 + \alpha_2 \frac{1}{\sqrt{X_i}} + v_i$
2	$ e_i  = \alpha_1 + \alpha_2 \sqrt{X_i} + v_i$	5	$ e_i  = \sqrt{\alpha_1 + \alpha_2 X_i} + v_i$
3	$ e_i  = \alpha_1 + \alpha_2 \frac{1}{X_i} + v_i$	6	$ e_i  = \sqrt{\alpha_1 + \alpha_2 X_i^2} + v_i$

Tiêu chuẩn kiểm định:  $\chi_{qs}^2 = nR^2$

Với  $R^2$  là hệ số xác định của mô hình hồi quy phụ.

Nếu  $\chi_{qs}^2 > \chi_\alpha^2(k-1)$  thì bác bỏ  $H_0$

### 5.1.1.2. Kiểm định Breusch – Pagan – Goldfrey (BPG, 1979)

Mô hình hồi quy phụ để kiểm định phương sai sai số thay đổi của BPG là:

$$e_i^2 = \alpha_1 + \alpha_2 X_{2i} + \dots + \alpha_k X_{ki} + v_i$$

Ta cần kiểm định cặp giả thiết sau:

$H_0: \alpha_2 = \dots = \alpha_k = 0 \rightarrow$  mô hình (6.1) có PSSS bằng nhau

$H_1: \exists \alpha_j \neq 0, j = \overline{2, k} \rightarrow$  mô hình (6.1) có PSSS thay đổi

Tiêu chuẩn kiểm định: (gọi hệ số xác định trong mô hình phụ là  $R^2$ )

Dùng  $\chi_{qs}^2 = nR^2$  (hoặc  $F_{qs} = \frac{R^2 / (k-1)}{(1-R^2) / (n-k)}$ )

Nếu:  $\chi_{qs}^2 > \chi_\alpha^2(k-1)$  (hoặc  $F_\alpha(k-1, n-k)$ ) thì bác bỏ  $H_0$

### 5.1.1.3. Kiểm định White

Dùng cho mô hình hồi quy đa biến và hơn nữa kiểm định này không đòi hỏi nhiều trong mô hình ban đầu phải có phân phối chuẩn.

White đề nghị hồi quy bình phương phần dư theo các biến độc lập, bình phương của các biến này và các tích chéo của biến độc lập.

Chẳng hạn, xét mô hình hồi quy 3 biến (tổng quát k biến là tương tự):

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$$

Mô hình hồi quy phụ của White có hai dạng như sau:

White – no cross (không có hệ số chéo, không có tích chéo):

$$e_i^2 = \alpha_1 + \alpha_2 X_{2i}^2 + \alpha_3 X_{3i}^2 + v_i \quad (a)$$

và White – cross (có hệ số chéo, có tích chéo):

$$e_i^2 = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{3i} + \alpha_4 X_{2i}^2 + \alpha_5 X_{3i}^2 + \alpha_6 X_{2i} X_{3i} + v_i \quad (b)$$

Mô hình phụ của White nhất thiết phải có hệ số chặn  $\alpha_1$  và gọi hệ số xác định trong các mô hình phụ là  $R^2$ . Trong mô hình phụ (b) thì  $\alpha_6$  được gọi là hệ số chéo và tích giữa hai biến độc lập  $X_2 X_3$  được gọi là tích chéo.

Cặp giả thiết cần kiểm định:

+ Đối với mô hình (a):

$H_0: \alpha_2 = \alpha_3 = 0$  (PSSS không thay đổi)

$H_1: \exists \alpha_j \neq 0, j = 2, 3$  (PSSS thay đổi)

+ Đối với mô hình (b):

$$H_0: \alpha_2 = \dots = \alpha_6 = 0 \quad (\text{PSSS không thay đổi})$$

$$H_1: \exists \alpha_j \neq 0, j = \overline{2,6} \quad (\text{PSSS thay đổi})$$

Tiêu chuẩn kiểm định  $\chi^2: \chi_{qs} = nR^2$ ; nếu  $\chi_{qs} > \chi_{\alpha}^2(k_* - 1)$  thì bác bỏ  $H_0$

### 5.1.2. Tính chuẩn của sai số ngẫu nhiên (SSNN)

#### Kiểm định Jarque – Bera (JB)

Cặp giả thiết

$H_0$  : SSNN phân phối chuẩn

$H_1$  : SSNN không phân phối chuẩn.

Tiêu chuẩn kiểm định

$$JB = n \left[ \frac{S^2}{6} + \frac{(K-3)^2}{24} \right]$$

trong đó S là hệ số bất đối xứng (Skewness) và K là hệ số nhọn (Kurtosis) được tính bởi công thức sau:

$$S = \frac{\sum e_i^3 / n}{(\sum e_i^2 / n)^{3/2}} \quad \text{và} \quad K = \frac{\sum e_i^4 / n}{(\sum e_i^2 / n)^2}$$

Nếu  $JB > \chi_{\alpha}^2(2)$  thì bác bỏ  $H_0$

### 5.1.3. Đa cộng tuyến

Xét mô hình hồi quy có từ hai biến độc lập trở lên ( $k \geq 3$ ):

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + u_i$$

**Giả thiết 4** : Giữa các biến độc lập không có hiện tượng đa cộng tuyến.

Nếu tồn tại mối quan hệ tuyến tính giữa các biến độc lập  $X_2, X_3, \dots, X_k$  thì mô hình hồi quy có khuyết tật đa cộng tuyến.

$X_2, X_3, \dots, X_k$  được gọi là có quan hệ đa cộng tuyến nếu  $\exists \lambda_2; \lambda_3; \dots; \lambda_k$  không đồng thời bằng 0 sao cho:

$$\left[ \begin{array}{ll} \lambda_2 X_{2i} + \dots + \lambda_k X_{ki} = 0 & \forall i \text{ đa cộng tuyến hoàn hảo} \\ & (\text{perfect multicollinearity}) \\ \lambda_2 X_{2i} + \dots + \lambda_k X_{ki} + v_i = 0 & \forall i \text{ đa cộng tuyến không hoàn hảo} \\ & (\text{imperfect multicollinearity}) \end{array} \right.$$

Trong thực tế đa cộng tuyến hoàn hảo thường không xảy ra, tuy nhiên đa cộng tuyến không hoàn hảo thì lại thường gặp. Nó tồn tại trong hầu hết các mô hình hồi quy có từ 2 biến độc lập trở lên.

#### Cách phát hiện đa cộng tuyến:

- Xem xét hệ số xác định của các mô hình hồi quy phụ.
- Nhân tử phóng đại phương sai (variance inflation factor)

Một thước đo quan trọng và thông dụng của đa cộng tuyến là nhân tử phóng đại phương sai VIF, gắn với biến  $X_j$ , kí hiệu là  $VIF(X_j)$ :

$$VIF(X_j) = \frac{1}{1 - R_j^2}$$

- Hệ số tương quan giữa các cặp biến độc lập cao

Phương pháp này cho biết khi hệ số tương quan cặp  $r_{ij}$  giữa các biến độc lập lớn hơn 0,8 thì đa cộng tuyến là mạnh.

## 5.2. Hướng dẫn Eviews

### 5.2.1. Kiểm định Glejer

Cho ví dụ sau đây. Tập số liệu CH5VD1.wfl với các biến số

Y: chi tiêu cho thực phẩm (USD), X là thu nhập của hộ gia đình (USD). Khảo sát 38 hộ trong một địa phương. Kết quả hồi quy:

Dependent Variable: Y  
Method: Least Squares  
Date: 04/24/14 Time: 11:07  
Sample: 1 38  
Included observations: 38

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X	0.157511	0.044590	3.532449	0.0011
C	6328.092	2749.312	2.301700	0.0272

R-squared	0.257398	Mean dependent var	15532.24
Adjusted R-squared	0.236770	S.D. dependent var	6189.981
S.E. of regression	5407.754	Akaike info criterion	20.08025
Sum squared resid	1.05E+09	Schwarz criterion	20.16644
Log likelihood	-379.5248	Hannan-Quinn criter.	20.11092
F-statistic	12.47820	Durbin-Watson stat	1.658798
Prob(F-statistic)	0.001150		

Kiểm tra hiện tượng phương sai sai số thay đổi cho mô hình này bằng Glejer như sau:

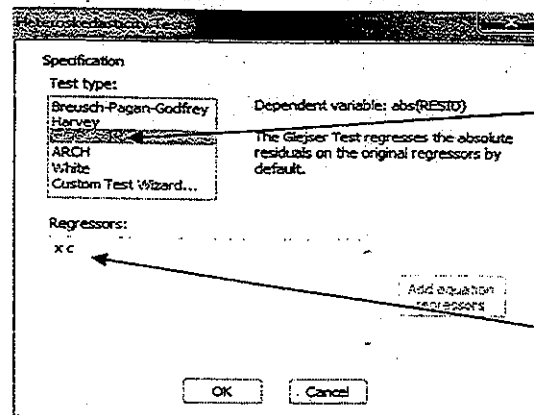
Cặp giả thiết cần kiểm định:

$H_0$  : mô hình không có khuyết tật PSSSTD

$H_1$  : mô hình có khuyết tật PSSSTD

[Equation] View → Stability Tests → Heteroskedasticity Test ...

Hộp thoại Heteroskedasticity Test



Chọn Glejer trong Test Type.

Ở đây còn khá nhiều lựa chọn khác để kiểm tra hiện tượng PSSSTD như: BPG, White ...

Điều chỉnh mô hình phụ cho Glejer ở đây. Mặc định biến phụ thuộc luôn là abs(Resid)

Nhấn OK ta được kết quả sau:

#### Heteroskedasticity Test: Glejser

F-statistic	0.981604	Prob. F(1,36)	0.3284
Obs*R-squared	1.008636	Prob. Chi-Square(1)	0.3152
Scaled explained SS	1.116471	Prob. Chi-Square(1)	0.2907

#### Test Equation:

Dependent Variable: ARESID  
Method: Least Squares  
Date: 04/24/14 Time: 14:25  
Sample: 1 38  
Included observations: 38

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2333.102	1767.503	1.319999	0.1952
X	0.028401	0.028666	0.990760	0.3284

R-squared	0.026543	Mean dependent var	3992.734
Adjusted R-squared	-0.000497	S.D. dependent var	3475.722
S.E. of regression	3476.586	Akaike info criterion	19.19669
Sum squared resid	4.35E+08	Schwarz criterion	19.28287
Log likelihood	-362.7370	Hannan-Quinn criter.	19.22735
F-statistic	0.981604	Durbin-Watson stat	1.596389
Prob(F-statistic)	0.328416		

Bảng kết quả kiểm định Glejer được chia làm hai thành phần. Phần trên là tóm tắt kết quả của các thống kê dùng trong

kiểm định Glejer (ta nên quan tâm đến thống kê Chi-square). Phần dưới là mô tả cụ thể về mô hình hồi quy phụ được sử dụng.

Theo như kết quả của Glejer cho ví dụ trên thì thống kê Chi-square = 1,0086 với p\_value = 0,3152 > 0,05 cho nên ta chưa đủ cơ sở bác bỏ  $H_0$ .

Vậy với mức ý nghĩa 5%, có thể cho rằng mô hình ban đầu không có hiện tượng phương sai số thay đổi.

### 5.2.2. Kiểm định White và BPG

Lấy lại tập số liệu CH4VD2.wfl nhưng chỉ xem xét sự phụ thuộc của giá bán nhà vào diện tích và giá vàng ta được:

Dependent Variable: GN  
Method: Least Squares  
Date: 04/24/14 Time: 14:40  
Sample: 1 100  
Included observations: 100

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DT	1.525973	0.297188	5.134710	0.0000
GV	17.50441	5.823651	3.005746	0.0034
C	-281.0716	256.9140	-1.094030	0.2767
R-squared	0.271333	Mean dependent var	639.7300	
Adjusted R-squared	0.256309	S.D. dependent var	207.2318	
S.E. of regression	178.7116	Akaike info criterion	13.23896	
Sum squared resid	3097969.	Schwarz criterion	13.31712	
Log likelihood	-658.9482	Hannan-Quinn criter.	13.27059	
F-statistic	18.05992	Durbin-Watson stat	0.406489	
Prob(F-statistic)	0.000000			

Đây là mô hình hồi quy bội nên để kiểm tra PSSS thay đổi ta dùng kiểm định White hoặc BPG.

Cặp giả thiết cần kiểm định:

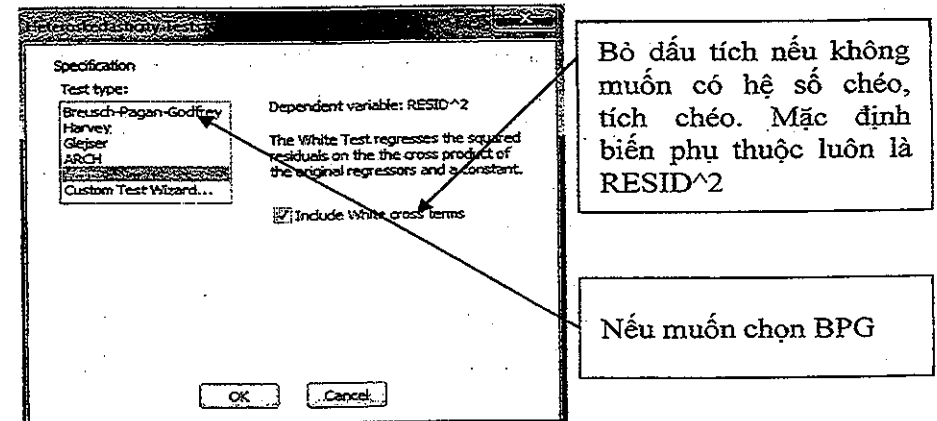
$H_0$ : mô hình không có khuyết tật PSSS thay đổi

$H_1$ : mô hình có khuyết tật PSSS thay đổi

Trong Eviews ta làm như sau:

[Equation] View → Residual Tests → Heteroskedasticity Tests ...

Hộp thoại Heteroskedasticity Test



Kết quả kiểm định như sau:

Heteroskedasticity Test: White

F-statistic	2.684559	Prob. F(5,94)	0.0259
Obs*R-squared	12.49529	Prob. Chi-Square(5)	0.0286
Scaled explained SS	6.202380	Prob. Chi-Square(5)	0.2870

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 04/24/14 Time: 14:48

Sample: 1 100

Included observations: 100

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-173878.7	729184.7	-0.238456	0.8120
DT	1324.955	742.1238	1.785355	0.0774
DT^2	0.601722	0.544368	1.105359	0.2718
DT*GV	-33.65074	15.78795	-2.131419	0.0357

GV	8955.053	34826.94	0.257130	0.7976
GV^2	-93.59924	413.0106	-0.226627	0.8212
R-squared	0.124953	Mean dependent var	30979.69	
Adjusted R-squared	0.078408	S.D. dependent var	31982.22	
S.E. of regression	30702.80	Akaike info criterion	23.56022	
Sum squared resid	8.86E+10	Schwarz criterion	23.71653	
Log likelihood	-1172.011	Hannan-Quinn criter.	23.62348	
F-statistic	2.684559	Durbin-Watson stat	1.228050	
Prob(F-statistic)	0.025903			

Kết quả cho thấy giá trị thống kê Chi-square = 12,495 với p\_value = 0,0286 (<5%) nên ta bác bỏ H<sub>0</sub>. Vậy mô hình ban đầu có hiện tượng phương sai sai số thay đổi.

### 5.2.3. Ước lượng sai số chuẩn

Khi mô hình có hiện tượng PSSS thay đổi, các ước lượng OLS cho các hệ số vẫn là ước lượng không chệch, chỉ có phương sai của các hệ số ước lượng và hiệp phương sai giữa các hệ số ước lượng là chệch. Từ đó, White đề xuất phương pháp sai số chuẩn vững với tư tưởng như sau: vẫn sử dụng các hệ số ước lượng từ các phương pháp OLS, tuy nhiên phương sai các hệ số ước lượng thì được tính toán lại mà không sử dụng đến giải thiết PSSS thay đổi.

Để tính các sai số chuẩn theo White trong Eviews ta làm như sau:

Hồi quy mô hình gốc, tiếp theo lần lượt thực hiện

[Equation]Estimate → Options → Heteroskedasticity... → White → OK

Trở lại ví dụ trên với tập số liệu CH4VD2.wfl và cũng chỉ xem xét sự phụ thuộc của giá bán nhà vào diện tích và giá vàng.

### Kết quả hồi quy mô hình gốc:

Dependent Variable: GN  
Method: Least Squares  
Date: 12/26/14 Time: 18:00  
Sample: 1 100  
Included observations: 100

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DT	1.525973	0.297188	5.134710	0.0000
GV	17.50441	5.823651	3.005746	0.0034
C	-281.0716	256.9140	-1.094030	0.2767

R-squared	0.271333	Mean dependent var	639.7300
Adjusted R-squared	0.256309	S.D. dependent var	207.2318
S.E. of regression	178.7116	Akaike info criterion	13.23896
Sum squared resid	3097969.	Schwarz criterion	13.31712
Log likelihood	-658.9482	Hannan-Quinn criter.	13.27059
F-statistic	18.05992	Durbin-Watson stat	0.406489
Prob(F-statistic)	0.000000		

### Kết quả ước lượng lại sai số chuẩn:

Dependent Variable: GN  
Method: Least Squares  
Date: 12/26/14 Time: 18:01  
Sample: 1 100  
Included observations: 100

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DT	1.525973	0.370455	4.119191	0.0001
GV	17.50441	6.527311	2.681719	0.0086
C	-281.0716	284.4670	-0.988064	0.3256

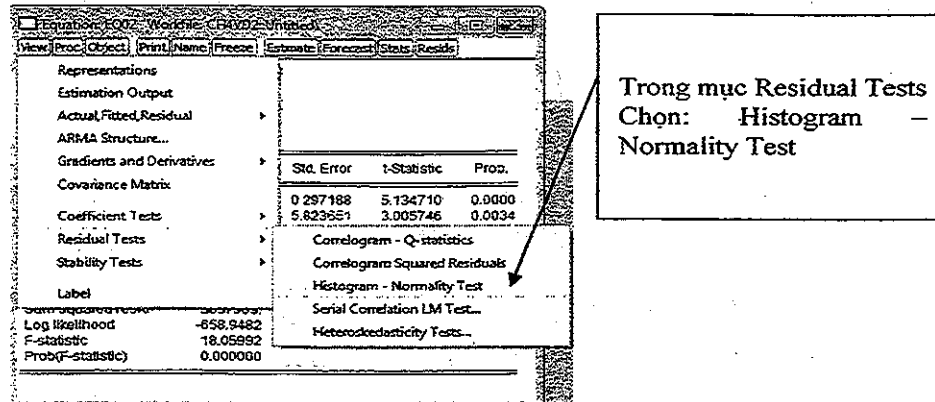
R-squared	0.271333	Mean dependent var	639.7300
Adjusted R-squared	0.256309	S.D. dependent var	207.2318
S.E. of regression	178.7116	Akaike info criterion	13.23896
Sum squared resid	3097969.	Schwarz criterion	13.31712
Log likelihood	-658.9482	Hannan-Quinn criter.	13.27059
F-statistic	18.05992	Durbin-Watson stat	0.406489
Prob(F-statistic)	0.000000		

So sánh hai bảng kết quả hồi quy ta thấy chỉ có sai số chuẩn  $se(\hat{\beta}_j)$  là thay đổi để đảm bảo chúng là các ước lượng vững.

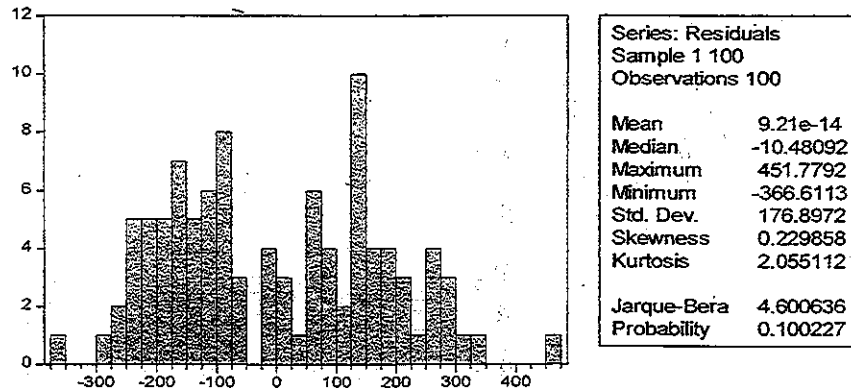
#### 5.2.4. Kiểm định tính chuẩn của sai số ngẫu nhiên

Thực hiện kiểm tra tính chuẩn cho nhiều của mô hình trên ta thực hiện lệnh:

[Equation] View → Residual Tests → Histogram – Normality Test



Kết quả kiểm định:



Sử dụng tiêu chuẩn Jarque Bera để kiểm tra tính chuẩn của sai số ngẫu nhiên thì cặp giả thiết là:

$H_0$  : sai số ngẫu nhiên có phân phối chuẩn

$H_1$  : sai số ngẫu nhiên không có phân phối chuẩn

Tiêu chuẩn Jarque – Bera = 4,6006 với p-value = 0,1002 > 5% nên chưa đủ cơ sở bác bỏ  $H_0$ . Vậy sai số ngẫu nhiên của mô hình có phân phối chuẩn.

#### 5.2.5. Kiểm định Ramsey – Reset

Mục đích của kiểm định này là xem có bỏ sót biến quan trọng trong mô hình hồi quy hay không (nhất là khi không có số liệu về biến bỏ sót đó). Nó còn gọi là kiểm định dạng hàm sai với cặp giả thiết:

$H_0$  : định dạng hàm đúng (không bỏ sót biến)

$H_1$  : định dạng hàm sai (bỏ sót biến)

Kiểm định này dựa trên ý tưởng sử dụng tổ hợp tuyến tính của các  $\hat{Y}^2, \hat{Y}^3, \hat{Y}^4, \dots$  thay cho biến bị bỏ sót với mô hình hồi quy phụ:

$$Y_i = \alpha_1 + \alpha_2 X_{2i} + \dots + \alpha_k X_{ki} + \gamma_1 \hat{Y}^2 + \gamma_2 \hat{Y}^3 + v_i$$

Bậc của  $\hat{Y}$  là tùy ý, thông thường ta chọn bậc 2 hoặc 3.

Giả thiết kiểm định:

$H_0: \gamma_1 = \gamma_2 = \gamma_3 = 0$  (không bỏ sót biến)

$H_1: \exists \gamma_j \neq 0, j = 1, 2, 3$  (bỏ sót biến)

Tiêu chuẩn kiểm định

$$F_{qs} = \frac{(R^2 - R_*^2)/m}{(1 - R_*^2)/(n - k)}$$

Với  $R^2, R_*^2$  tương ứng là hệ số xác định của mô hình gốc, mô hình phụ

$m$  là số biến độc lập mới được đưa vào (ở đây  $m = 2$ )

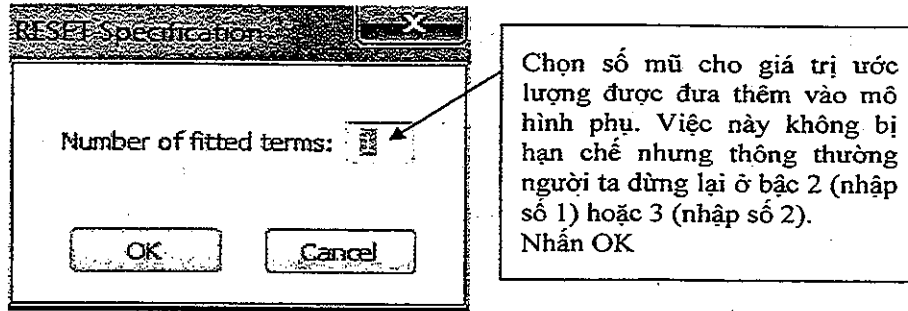
$k$  là số hệ số hồi quy của mô hình phụ.

Nếu  $F_{qs} > F_{\alpha}(m, n - k)$  thì bác bỏ giả thiết  $H_0$ .

Trong Eviews ta làm như sau:

[Equation] View → Stability Tests → Ramsey RESET Test ...

Hộp thoại RESET Specification



Tiếp tục với tập số liệu CH4VD2.wf1. Thực hiện các bước kiểm định ta nhận được kết quả như sau:

Ramsey RESET Test:

F-statistic	12.32895	Prob. F(2,95)	0.0000
Log likelihood ratio	23.07600	Prob. Chi-Square(2)	0.0000

Test Equation:

Dependent Variable: GN

Method: Least Squares

Date: 04/24/14 Time: 16:10

Sample: 1 100

Included observations: 100

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DT	22.56032	13.77319	1.637988	0.1047

GV	232.5267	155.3348	1.496940	0.1377
C	-6964.773	4441.851	-1.567989	0.1202
FITTED^2	-0.014631	0.013095	-1.117266	0.2667
FITTED^3	4.56E-06	6.18E-06	0.736993	0.4629

R-squared	0.421490	Mean dependent var	639.7300
Adjusted R-squared	0.397131	S.D. dependent var	207.2318
S.E. of regression	160.9044	Akaike info criterion	13.04820
Sum squared resid	2459570.	Schwarz criterion	13.17846
Log likelihood	-647.4102	Hannan-Quinn criter.	13.10092
F-statistic	17.30372	Durbin-Watson stat	0.575390
Prob(F-statistic)	0.000000		

Trong kết quả trên ta đang dùng tới bậc 3 của giá trị FITTED (tức là khi nhập trong ô Number of fitted terms là 2).

Thông kê  $F = 12,329$  với  $p\text{-value} = 0,0000 < 5\%$  cho nên ta bác bỏ  $H_0$ . Kết luận mô hình ban đầu định dạng hàm sai, có khả năng bỏ sót biến nào đó mà ta không biết chính xác.

### 5.3. Bài tập

**Bài 5.1.** Biết  $D$  là biến giả,  $X$  là biến định lượng nào đó. Hãy cho biết giữa  $D$  và  $DX (= D_i * X_i)$  có mối quan hệ đa cộng tuyến hay không?

**Bài 5.2.** Klein và Golberger dùng mô hình sau đây cho số liệu của nước Mỹ, từ 1928 – 1952 (ngoại trừ các năm 1942 – 1944).

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + u_t$$

Trong đó :  $Y$  – tiêu dùng,  $X_2$  – thu nhập từ tiền lương,  $X_3$  – thu nhập ngoài tiền lương, ngoài trang trại,  $X_4$  – thu nhập từ trang trại.



Dependent Variable: Y

Included observations: 22 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X2	1.053033	0.164935	6.384550	0.0000
X3	-0.264098	0.309092	-0.854431	0.4041
X4	0.538580	1.022058	0.526957	0.6047
C	18.78809	3.485423	5.390477	0.0000
R-squared	0.961311	Mean dependent var	75.87273	
Adjusted R-squared	0.954863	S.D. dependent var	21.25673	
F-statistic	149.0818	Durbin-Watson stat	1.364497	
Prob(F-statistic)	0.000000			

1. Hãy tìm điểm không hợp lý về mặt kinh tế của mô hình trên.
2. Trong mô hình trên có dấu hiệu của hiện tượng đa cộng tuyến hay không?
3. Hồi quy  $X_2$  theo  $X_3$  và  $X_4$  ta thu được kết quả dưới đây. Hãy cho biết mức độ đa cộng tuyến trong mô hình ban đầu?

Dependent Variable: X2

Included observations: 22 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X3	1.466698	0.267619	5.480537	0.0000
X4	5.473395	0.666555	8.211467	0.0000
C	-2.393212	4.816857	-0.496841	0.6250
R-squared	0.909100	Mean dependent var	56.26091	
Adjusted R-squared	0.899531	S.D. dependent var	19.81804	
F-statistic	95.01002	Durbin-Watson stat	1.738604	

**Bài 5.3.** Một người muốn phân tích tình hình doanh thu của một bộ phim được phát hành với các biến là: chi phí sản xuất, chi phí quảng cáo và yếu tố sau:

- $D = 1$  nếu bộ phim đã được giới thiệu trên ít nhất một tạp chí trước khi phát hành
- $D = 0$  nếu ngược lại

Dependent Variable: DT

Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CPSX	3.189054	0.438662	7.269957	0.0000
CPQC	2.155770	0.291154	7.404238	0.0000
D	7.991656	2.118721	3.771925	0.0017
C	5.180920	2.700336	1.918620	0.0731
R-squared	0.954499	Mean dependent var	46.05000	
Adjusted R-squared	0.945967	S.D. dependent var	18.55710	
F-statistic	111.8795			
Prob(F-statistic)	0.000000			

1. Có thể có đa cộng tuyến trong mô hình hồi quy này hay không?
2. Người ta ước lượng tiếp mô hình sau. Có thể có khả năng xảy ra đa cộng tuyến của mô hình ban đầu hay không?

Dependent Variable: CPSX

Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CPQC	0.353570	0.136237	2.595260	0.0189
D	1.536332	1.110596	1.383339	0.1845
C	4.466279	1.027470	4.346871	0.0004
R-squared	0.399263	Mean dependent var	7.415000	
Adjusted R-squared	0.328588	S.D. dependent var	2.910647	
F-statistic	5.649289	Durbin-Watson stat	2.125171	
Prob(F-statistic)	0.013147			

**Bài 5.4.** Trong mô hình hồi quy có chứa các biến kinh tế có số liệu chuỗi thời gian như: GNP, lượng cung tiền, giá cả, thu nhập, tỷ lệ thất nghiệp ... thường có đa cộng tuyến. Tại sao?

**Bài 5.5.** Cho dữ liệu về GDP, Imports và CPI của Mỹ từ năm 1975 – 2005. Hồi quy mô hình sau đây:

Mô hình (1)

Dependent Variable: LOG(IMPORTS)  
Method: Least Squares  
Sample: 1975 2005

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(GDP)	1.850099	0.182912	10.11471	0.0000
LOG(CPI)	-0.873369	0.284805	-3.066548	0.0048
C	1.409416	0.270074	5.218629	0.0000
R-squared	0.992005	Mean dependent var	13.08472	
Adjusted R-squared	0.991434	S.D. dependent var	0.762092	
S.E. of regression	0.070532	Akaike info criterion	-2.373736	
Sum squared resid	0.139293	Schwarz criterion	-2.234963	
Log likelihood	39.79291	Hannan-Quinn criter.	-2.328500	
F-statistic	1737.193	Durbin-Watson stat	0.650260	
Prob(F-statistic)	0.000000			

1. Có dấu hiệu của đa cộng tuyến trong mô hình hồi quy này hay không?

2. Người ta ước lượng tiếp các mô hình sau:

Mô hình (2)

Dependent Variable: LOG(IMPORTS)  
Method: Least Squares  
Sample: 1975 2005

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(GDP)	1.293252	0.024951	51.83130	0.0000
C	2.001910	0.214308	9.341269	0.0000
R-squared	0.989321	Mean dependent var	13.08472	
Adjusted R-squared	0.988952	S.D. dependent var	0.762092	

S.E. of regression	0.080102	Akaike info criterion	-2.148687
Sum squared resid	0.186074	Schwarz criterion	-2.056171
Log likelihood	35.30465	Hannan-Quinn criter.	-2.118529
F-statistic	2686.484	Durbin-Watson stat	0.517695
Prob(F-statistic)	0.000000		

Mô hình (3)

Dependent Variable: LOG(IMPORTS)  
Method: Least Squares  
Sample: 1975 2005

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(CPI)	1.986499	0.072515	27.39449	0.0000
C	3.578293	0.348057	10.28076	0.0000
R-squared	0.962795	Mean dependent var	13.08472	
Adjusted R-squared	0.961512	S.D. dependent var	0.762092	
S.E. of regression	0.149510	Akaike info criterion	-0.900561	
Sum squared resid	0.648248	Schwarz criterion	-0.808045	
Log likelihood	15.95869	Hannan-Quinn criter.	-0.870403	
F-statistic	750.4582	Durbin-Watson stat	0.279792	
Prob(F-statistic)	0.000000			

Mô hình (4)

Dependent Variable: LOG(GDP)  
Method: Least Squares  
Sample: 1975 2005

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(CPI)	1.545793	0.034730	44.50942	0.0000
C	1.172304	0.166696	7.032591	0.0000
R-squared	0.985573	Mean dependent var	8.569723	
Adjusted R-squared	0.985075	S.D. dependent var	0.586128	
S.E. of regression	0.071605	Akaike info criterion	-2.372953	
Sum squared resid	0.148693	Schwarz criterion	-2.280437	
Log likelihood	38.78076	Hannan-Quinn criter.	-2.342795	
F-statistic	1981.089	Durbin-Watson stat	0.199406	
Prob(F-statistic)	0.000000			

Dựa vào các mô hình vừa hồi quy, hãy cho biết bản chất trong đa cộng tuyến của mô hình ban đầu?

**Bài 5.6.** Cho các biến như sau:

- BUSTRAVL: Mức độ giao thông bằng xe buýt ở đô thị tính theo ngàn hành khách mỗi giờ
- FARE: Giá vé xe buýt tính bằng USD
- GASPRICE: Giá một ga lông nhiên liệu tính bằng USD
- INCOME: Thu nhập bình quân đầu người tính bằng USD
- POP: Dân số thành phố tính bằng ngàn người
- DENSITY: Mật độ dân số tính (người/dặm vuông)
- LANDAREA: Diện tích thành phố (dặm vuông)

Số liệu lấy cho 40 thành phố ở nước Mỹ.

Dependent Variable: BUSTRAVL  
Method: Least Squares  
Sample: 1 40  
Included observations: 40

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DENSITY	0.116415	0.059570	1.954253	0.0592
FARE	-238.6544	451.7281	-0.528314	0.6008
GASPRICE	522.1132	2658.228	0.196414	0.8455
INCOME	-0.194744	0.064887	-3.001294	0.0051
LANDAREA	-1.155230	1.802638	-0.640855	0.5260
POP	1.711442	0.231364	7.397176	0.0000
C	2744.680	2641.672	1.038994	0.3064

R-squared	0.921026	Mean dependent var	1933.175
Adjusted R-squared	0.906667	S.D. dependent var	2431.757
S.E. of regression	742.9113	Akaike info criterion	16.21666
Sum squared resid	18213267	Schwarz criterion	16.51221
Log likelihood	-317.3332	Hannan-Quinn criter.	16.32352
F-statistic	64.14338	Durbin-Watson stat	2.082671
Prob(F-statistic)	0.000000		

Thực hiện kiểm định BPG và White cho mô hình trên ta được:

**Heteroskedasticity Test: Breusch-Pagan-Godfrey**

F-statistic	2.238110	Prob. F(6,33)	0.0638
Obs*R-squared	11.56928	Prob. Chi-Square(6)	0.0723
Scaled explained SS	8.471470	Prob. Chi-Square(6)	0.2056

Test Equation:  
Dependent Variable: RESID^2  
Method: Least Squares  
Date: 05/22/13 Time: 00:31  
Sample: 1 40  
Included observations: 40

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-442974.3	2204417.	-0.200948	0.8420
DENSITY	-29.88672	49.71010	-0.601220	0.5518
FARE	-878904.3	376957.3	-2.331575	0.0260
GASPRICE	2280800.	2218233.	1.028206	0.3113
INCOME	-30.73957	54.14658	-0.567710	0.5741
LANDAREA	654.8321	1504.262	0.435318	0.6662
POP	224.0891	193.0685	1.160671	0.2541

R-squared	0.289232	Mean dependent var	455331.7
Adjusted R-squared	0.160002	S.D. dependent var	676414.3
S.E. of regression	619943.3	Akaike info criterion	29.67027
Sum squared resid	1.27E+13	Schwarz criterion	29.96583
Log likelihood	-586.4054	Hannan-Quinn criter.	29.77713
F-statistic	2.238110	Durbin-Watson stat	1.887274
Prob(F-statistic)	0.063848		

**Heteroskedasticity Test: White**

F-statistic	2.240704	Prob. F(6,33)	0.0636
Obs*R-squared	11.57881	Prob. Chi-Square(6)	0.0721
Scaled explained SS	8.478447	Prob. Chi-Square(6)	0.2051

Test Equation:  
Dependent Variable: RESID^2  
Method: Least Squares  
Date: 05/22/13 Time: 00:32  
Sample: 1 40  
Included observations: 40

Variable	Coefficient	Std. Error	t-Statistic	Prob.
----------	-------------	------------	-------------	-------

C	-303266.6	1070316.	-0.283343	0.7787
DENSITY^2	-0.001914	0.002267	-0.844560	0.4044
FARE^2	-513333.3	194951.2	-2.633138	0.0128
GASPRICE^2	1515218.	1162251.	1.303693	0.2014
INCOME^2	-0.000501	0.001613	-0.310775	0.7579
LANDAREA^2	3.176416	1.993660	1.593258	0.1206
POP^2	0.029530	0.025602	1.153445	0.2570
R-squared	0.289470	Mean dependent var	455331.7	
Adjusted R-squared	0.160283	S.D. dependent var	676414.3	
S.E. of regression	619839.4	Akaike info criterion	29.66994	
Sum squared resid	1.27E+13	Schwarz criterion	29.96549	
Log likelihood	-586.3987	Hannan-Quinn criter.	29.77680	
F-statistic	2.240704	Durbin-Watson stat	1.973689	
Prob(F-statistic)	0.063581			

Heteroskedasticity Test: White

F-statistic	2.498772	Prob. F(27,12)	0.0489
Obs*R-squared	33.95974	Prob. Chi-Square(27)	0.1672
Scaled explained SS	24.86662	Prob. Chi-Square(27)	0.5819

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 05/22/13 Time: 00:33

Sample: 1 40

Included observations: 40

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-67226492	49111463	-1.368855	0.1961
DENSITY	-1861.429	2219.391	-0.838711	0.4180
DENSITY^2	0.040462	0.046976	0.861329	0.4059
DENSITY*FARE	-661.8972	434.7548	-1.522461	0.1538
DENSITY*GASPRICE	2117.683	1930.091	1.097193	0.2941
DENSITY*INCOME	-0.006942	0.056184	-0.123563	0.9037
DENSITY*LANDAREA	5.148586	6.796126	0.757577	0.4633
DENSITY*POP	-0.724115	0.777817	-0.930958	0.3702
FARE	9087177.	18641297	0.487476	0.6347
FARE^2	-2323641.	3345294.	-0.694600	0.5005
FARE*GASPRICE	-15423587	17004422	-0.907034	0.3822
FARE*INCOME	852.8108	442.8861	1.925576	0.0782
FARE*LANDAREA	-22178.24	18422.74	-1.203851	0.2519
FARE*POP	3162.818	2661.092	1.188541	0.2576
GASPRICE	97731747	73069641	1.337515	0.2059

GASPRICE^2	-18035234	41295993	-0.436731	0.6701
GASPRICE*INCOME	-3851.236	3066.112	-1.256065	0.2330
GASPRICE*LANDAREA				
A	44891.53	64820.16	0.692555	0.5018
GASPRICE*POP	-11685.43	14008.79	-0.834150	0.4205
INCOME	2841.048	3394.521	0.836951	0.4190
INCOME^2	0.001292	0.027657	0.046716	0.9635
INCOME*LANDAREA	0.966437	4.180340	0.231186	0.8211
INCOME*POP	-0.373745	1.212276	-0.308300	0.7631
LANDAREA	-69211.68	112133.9	-0.617224	0.5486
LANDAREA^2	75.96018	76.79108	0.989180	0.3421
LANDAREA*POP	-24.62556	27.83978	-0.884546	0.3938
POP	19731.16	25656.07	0.769064	0.4567
POP^2	2.048598	2.462429	0.831942	0.4217

R-squared	0.848994	Mean dependent var	455331.7
Adjusted R-squared	0.509229	S.D. dependent var	676414.3
S.E. of regression	473862.2	Akaike info criterion	29.17125
Sum squared resid	2.69E+12	Schwarz criterion	30.35346
Log likelihood	-555.4250	Hannan-Quinn criter.	29.59870
F-statistic	2.498772	Durbin-Watson stat	1.717110
Prob(F-statistic)	0.048938		

Có kết luận gì khuyết tật phương sai thay đổi của mô hình hồi quy ban đầu?

Bài 5.7. Lấy bộ số liệu của bài 5.6. Xét mô hình hồi quy sau:

Dependent Variable: BUSTRAVL

Method: Least Squares

Date: 05/22/13 Time: 00:36

Sample: 1 40

Included observations: 40

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GASPRICE	17955.06	7119.116	2.522091	0.0160
C	-14482.24	6518.628	-2.221670	0.0323
R-squared	0.143391	Mean dependent var	1933.175	
Adjusted R-squared	0.120848	S.D. dependent var	2431.757	
S.E. of regression	2280.091	Akaike info criterion	18.35052	
Sum squared resid	1.98E+08	Schwarz criterion	18.43497	
Log likelihood	-365.0105	Hannan-Quinn criter.	18.38106	
F-statistic	6.360945	Durbin-Watson stat	1.965217	
Prob(F-statistic)	0.015976			

Thực hiện kiểm định Glejer cho mô hình trên ta được:

Heteroskedasticity Test: Glejser

F-statistic	9.875458	Prob. F(1,38)	0.0032
Obs*R-squared	8.250957	Prob. Chi-Square(1)	0.0041
Scaled explained SS	12.09164	Prob. Chi-Square(1)	0.0005

Test Equation:

Dependent Variable: ARESID

Method: Least Squares

Date: 05/22/13 Time: 00:37

Sample: 1 40

Included observations: 40

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-5327.454	2177.398	-2.446706	0.0192
GASPRICE^2	8111.315	2581.146	3.142524	0.0032

R-squared	0.206274	Mean dependent var	1473.214
Adjusted R-squared	0.185386	S.D. dependent var	1685.081
S.E. of regression	1520.886	Akaike info criterion	17.54068
Sum squared resid	87897599	Schwarz criterion	17.62512
Log likelihood	-348.8136	Hannan-Quinn criter.	17.57121
F-statistic	9.875458	Durbin-Watson stat	1.874747
Prob(F-statistic)	0.003242		

Và kiểm định sau:

Dependent Variable: LOG(RESID^2)

Method: Least Squares

Date: 05/22/13 Time: 00:39

Sample: 1 40

Included observations: 40

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(GASPRICE)	14.31942	5.905228	2.424871	0.0202
C	14.96984	0.629869	23.76660	0.0000

R-squared	0.134002	Mean dependent var	13.66416
Adjusted R-squared	0.111212	S.D. dependent var	2.192411
S.E. of regression	2.066907	Akaike info criterion	4.338691
Sum squared resid	162.3400	Schwarz criterion	4.423135
Log likelihood	-84.77381	Hannan-Quinn criter.	4.369223
F-statistic	5.880001	Durbin-Watson stat	2.026363
Prob(F-statistic)	0.020176		

Hãy cho biết kết quả nhận được từ hai kiểm định trên là gì?

Bài 5.8. Cho bộ số liệu với các biến số:

**loans** = commercial and industrial loans (billions of dollars, not seasonally adjusted) khoản vay cho công nghiệp và thương nghiệp (tỷ USD)

**deposits** = tổng mức tín dụng trong hệ thống ngân hàng (tỷ USD)

**Tbill3** = lợi suất trái phiếu kho bạc kỳ hạn 3 tháng (%)

**prime** = lãi suất cơ bản bình quân (%)

**aaa** = lãi lợi suất đáo hạn bình quân của trái phiếu dài hạn do Moody's Investor Service (%)

**indprod** = index of industrial production, chỉ số sản xuất công nghiệp

Số liệu theo tháng từ tháng 01/1973 - 09/2003

Nguồn: Federal Reserve Board (<http://www.federalreserve.gov/releases/>)

1. Ước lượng các mô hình hồi quy tuyến tính thông thường, tuyến tính log, semi - log.
2. Các mô hình trên có đa cộng tuyến không?
3. Kiểm định hiện tượng phương sai sai số thay đổi cho mô hình trên?
4. Nhiễu của các mô hình vừa làm có tính chuẩn không?
5. Thực hiện Ramsey - Reset cho các mô hình và kết luận
6. Nếu có khuyết tật, hãy khắc phục để có kết quả ước lượng tốt nhất.

## Chương 6: MÔ HÌNH CHUỖI THỜI GIAN

### 6.1. Mô hình với số liệu chuỗi thời gian

#### 6.1.1. Mô hình

Một số mô hình hồi quy chuỗi thời gian cơ bản:

+ Mô hình hồi quy tĩnh:

$$Y_t = \beta_1 + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + u_t$$

Ở đây chúng ta chỉ xét đến quan hệ giữa các biến số tĩnh tại cùng một thời điểm. Mô hình này cho phép xem xét mối quan hệ tức thời hoặc mối quan hệ cân bằng dài hạn giữa các biến số.

+ Mô hình hồi quy động:

*Mô hình có trễ phân phối hữu hạn*

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \dots + \beta_p X_{t-p} + u_t$$

Mô hình này cho rằng tác động của sự thay đổi trong biến  $X$  lên biến  $Y$  sẽ triệt tiêu sau  $p$  thời kỳ. Nó được sử dụng rộng rãi khi hồi quy số liệu chuỗi thời gian do tác động trễ giữa các biến số kinh tế là phổ biến.

*Mô hình tự hồi quy*

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \dots + \beta_p Y_{t-p} + u_t$$

Đây là mô hình tự hồi quy bậc  $p$ , ký hiệu là AR( $p$ ). Mô hình này cũng có thể chứa biến khác như thông thường.

*Mô hình hồi quy với xu thế thời gian*

- Xu thế tuyến tính:  $X_t = a + bt + e_t$

Trong đó  $a$  và  $b$  là các tham số,  $e_t$  là sai số ngẫu nhiên thể

hiện tác động của các yếu tố khác lên  $X$

- Xu thế bậc hai:  $X_t = a + bt + ct^2 + e_t$
- Xu thế bậc ba:  $X_t = a + bt + ct^2 + dt^3 + e_t$
- Xu thế dạng mũ:  $\ln X_t = a + bt + e_t$

Ngoài ra, chúng ta còn có những dạng mô hình với yếu tố mùa vụ, mô hình chuỗi thời gian thay đổi cấu trúc đã được nhắc đến ở chương 4.

#### 6.1.2. Hiện tượng tự tương quan

Tự tương quan có thể hiểu là sự tương quan giữa các thành phần của chuỗi các quan sát được sắp xếp theo thứ tự thời gian (trong các số liệu chuỗi thời gian) hoặc không gian (trong số liệu chéo).

Xét mô hình hồi quy ban đầu:

$$Y_t = \beta_1 + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + u_t$$

Khi mô hình trên có hiện tượng tự tương quan, nghĩa là sai số ngẫu nhiên  $u$  tại các thời điểm khác nhau là có tương quan với nhau.

*Tự tương quan bậc 1*

$$u_t = \rho u_{t-1} + \epsilon_t$$

- +  $\rho < 0$ : mô hình có tự tương quan bậc 1 âm.
- +  $\rho > 0$ : mô hình có tự tương quan bậc 1 dương.
- +  $\rho = 0$ : mô hình không có tự tương quan.

*Tự tương quan bậc  $p$*

$$u_t = \rho_1 u_{t-1} + \dots + \rho_p u_{t-p} + \epsilon_t$$

##### 6.1.2.1. Kiểm định Durbin - Watson

Xét mô hình:  $Y_t = \beta_1 + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + u_t$

Kiểm định Durbin-Watson chỉ dùng để kiểm định tự tương quan bậc nhất:

$$u_t = \rho u_{t-1} + \epsilon_t \quad (7.1)$$

Thống kê Durbin - Watson: 
$$d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}$$

Thống kê  $d$  có thể biểu diễn xấp xỉ bởi công thức:

$$d \approx 2(1 - \hat{\rho})$$

Trong đó  $\hat{\rho}$  là ước lượng của  $\rho$  trong (7.1) với  $0 \leq \hat{\rho} \leq 1$ , do đó

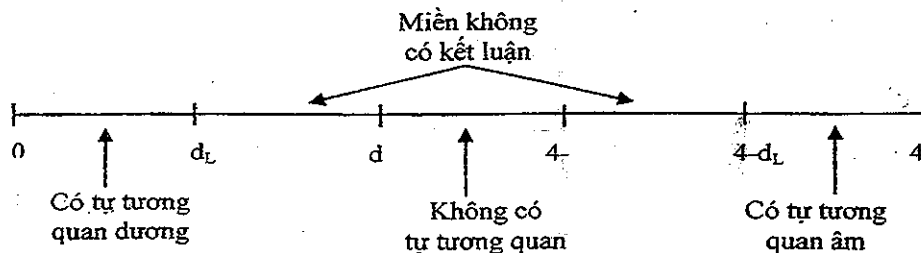
$$0 \leq d \leq 4$$

Ta xét các trường hợp đặc biệt của  $d$ :

- $d = 0 \Rightarrow \hat{\rho} = 1$ : dấu hiệu của tự tương quan dương rất cao giữa  $u_t$  và  $u_{t-1}$ .
- $d = 4 \Rightarrow \hat{\rho} = -1$ : dấu hiệu của tự tương quan âm rất cao giữa  $u_t$  và  $u_{t-1}$ .
- $d = 2 \Rightarrow \hat{\rho} = 0$ : dấu hiệu của không có tự tương quan dương hoặc tự tương quan âm giữa  $u_t$  và  $u_{t-1}$ .

Dựa trên kích thước mẫu  $n$ , số biến độc lập trong mô hình  $k' = k - 1$  và mức ý nghĩa  $\alpha$  cho trước, tra bảng tìm  $d_L$  và  $d_U$ .

Quy tắc kiểm định của Durbin - Watson:



**Lưu ý:** + Kiểm định này không áp dụng được với mô hình có biến giải thích là biến trễ của biến phụ thuộc.

+ Không có quan sát bị mất ở giữa.

+ Mô hình phải có hệ số chặn.

### Kiểm định Durbin-h

Kiểm định này để khắc phục trường hợp mô hình có chứa biến giải thích là biến trễ của biến phụ thuộc. Chẳng hạn xét mô hình hồi quy sau:

$$Y_t = \beta_1 + \beta_2 Y_{t-1} + u_t \quad (7.2)$$

Khi đó, kiểm định Durbin-h sử dụng thống kê sau đây:

$$h = \left(1 - \frac{d}{2}\right) \sqrt{\frac{n}{1 - n \cdot \text{var}(\hat{\beta}_2)}}$$

với  $d$  là thống kê Durbin-Watson của mô hình (7.2).

Việc đưa ra kết luận trong kiểm định này dựa trên sự so sánh  $h$  với giá trị tới hạn của quy luật chuẩn hóa.

### 6.1.2.2. Kiểm định Breusch - Godfrey

Xét mô hình hồi quy:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i$$

Để kiểm định có tự tương quan bậc  $p$  trong mô hình này hay không BG sử dụng cặp giả thiết sau:

$$H_0 : \rho_1 = \rho_2 = \dots = \rho_p = 0 \rightarrow \text{không có tự tương quan bậc } p$$

$$H_1 : \rho_1^2 + \rho_2^2 + \dots + \rho_p^2 \neq 0 \rightarrow \text{có tự tương quan ở bậc tương ứng}$$

Có 2 cách kiểm định cho cặp giả thiết này:

**Tiêu chuẩn kiểm định**  $\chi_{qs}^2$

Mô hình hồi quy phụ:

$$e_t = \beta_1 + \beta_2 X_t + \dots + \beta_k X_{kt} + \alpha_1 e_{t-1} + \alpha_2 e_{t-2} + \dots + \alpha_p e_{t-p} + v_t \quad (*)$$

Tiêu chuẩn kiểm định  $\chi^2: \chi_{qs}^2 = n R_{qs}^2 = (n-p) R_{qs}^2$  ;

Nếu  $\chi_{qs}^2 > \chi_{\alpha}^2(p)$  thì bác bỏ  $H_0$

**Tiêu chuẩn kiểm định  $F_{qs}$**

Ta thu hẹp mô hình hồi quy phụ (\*) thành mô hình sau:

$$e_t = \beta_1 + \beta_2 X_t + \dots + \beta_k X_{kt} + v_t \quad (**)$$

Tiêu chuẩn kiểm định :  $F_{qs} = \frac{(R_{qs}^2 - R_{..}^2) / p}{(1 - R_{qs}^2) / (n - k)}$

Nếu  $F_{qs} > F_{\alpha}(p; n-k)$  thì bác bỏ  $H_0$

## 6.2. Hướng dẫn Eviews

Lấy lại tập số liệu **CH4VD1.wfl** với các biến số sau đây:

PCE – PCE services (chi tiêu tiêu dùng cá nhân cho dịch vụ) (tỷ USD)

GDP – Gross domestic product (tổng sản phẩm quốc dân) (tỷ USD)

FFR - Federal Funds Rate (lãi suất quỹ dự trữ của Úc) (%)

M2 – lượng cung tiền (tỷ USD)

Kết quả ước lượng cho nền kinh tế Úc từ quý 1/1983 đến quý 1/2008

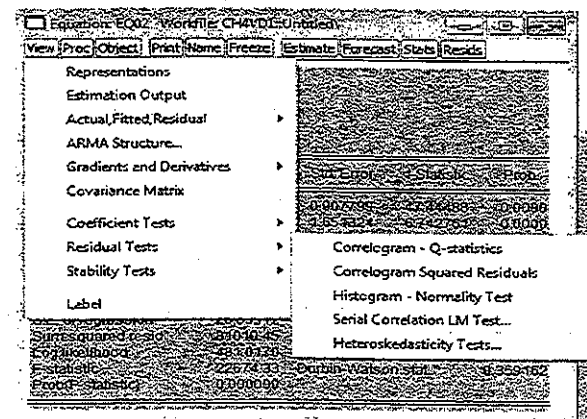
Dependent Variable: PCE  
Method: Least Squares  
Date: 12/27/14 Time: 15:38  
Sample: 1983Q1 2008Q1  
Included observations: 101

Variable	Coefficient	Std. Error	t-Statistic	Prob.
FFR	-11.13448	1.651324	-6.742761	0.0000

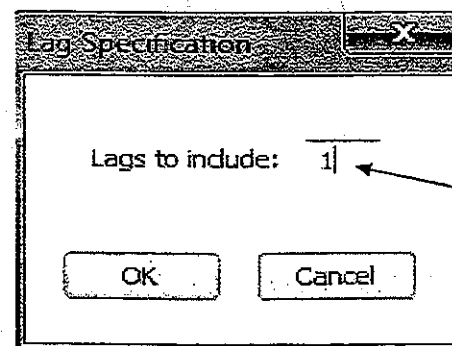
GDP	0.370003	0.007799	47.44488	0.0000
M2	0.033806	0.009395	3.598469	0.0005
C	194.9818	35.11981	5.551904	0.0000
R-squared	0.998576	Mean dependent var	3376.426	
Adjusted R-squared	0.998532	S.D. dependent var	754.2616	
S.E. of regression	28.89912	Akaike info criterion	9.604298	
Sum squared resid	81010.45	Schwarz criterion	9.707867	
Log likelihood	-481.0170	Hannan-Quinn criter.	9.646226	
F-statistic	22674.33	Durbin-Watson stat	0.359162	

Thực hiện kiểm định Breusch – Godfrey bằng Eviews như sau:

[Equation] View → Residual Tests → Serial Correlation LM Test...



Hộp thoại RESET Specification



Chỉ định độ trễ, tức là bậc của tự tương quan. Ở đây ta kiểm tra tự tương quan bậc nhất nên gõ 1.  
Nhấn OK



### Kết quả nhận được:

#### Breusch-Godfrey Serial Correlation LM Test:

F-statistic	188.4865	Prob. F(1,96)	0.0000
Obs*R-squared	66.91754	Prob. Chi-Square(1)	0.0000

#### Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 04/24/14 Time: 23:03

Sample: 1983Q1 2008Q1

Included observations: 101

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP	0.001217	0.004555	0.267306	0.7898
FFR	1.832648	0.973441	1.882649	0.0628
M2	0.001008	0.005486	0.183721	0.8546
C	-24.16057	20.58261	-1.173834	0.2434
RESID(-1)	0.828160	0.060322	13.72904	0.0000

R-squared	0.662550	Mean dependent var	9.31E-14
Adjusted R-squared	0.648490	S.D. dependent var	28.46234
S.E. of regression	16.87484	Akaike info criterion	8.537762
Sum squared resid	27336.98	Schwarz criterion	8.667224
Log likelihood	-426.1570	Hannan-Quinn criter.	8.590172
F-statistic	47.12164	Durbin-Watson stat	1.965862
Prob(F-statistic)	0.000000		

Kiểm định BG có cặp giả thiết:

$H_0$  : mô hình ban đầu không có tự tương quan bậc nhất

$H_1$  : mô hình ban đầu có tự tương quan bậc nhất

Theo bảng Eviews trên ta thấy thống kê  $F = 188,4865$  với  $p\text{-value} = 0,0000 < 5\%$  nên ta bác bỏ  $H_0$ . Vậy mô hình hồi quy có hiện tượng tự tương quan bậc nhất.

**Khắc phục tự tương quan bằng phương pháp Corchanre Orcutt**

File **ch6.2** là số liệu về tiêu dùng của người dân (CONS) và thu nhập (GDP) từ năm 1991 đến năm 2005 của Việt Nam.

Dependent Variable: CONS

Method: Least Squares

Date: 05/10/14 Time: 00:25

Sample: 1991 2005

Included observations: 15

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP	0.527252	0.032631	16.15797	0.0000
C	64882.05	14347.45	4.522201	0.0006

R-squared	0.952569	Mean dependent var	265765.4
Adjusted R-squared	0.948920	S.D. dependent var	122717.9
S.E. of regression	27735.32	Akaike info criterion	23.42237
Sum squared resid	1.00E+10	Schwarz criterion	23.51677
Log likelihood	-173.6678	Hannan-Quinn criter.	23.42136
F-statistic	261.0800	Durbin-Watson stat	0.295616
Prob(F-statistic)	0.000000		

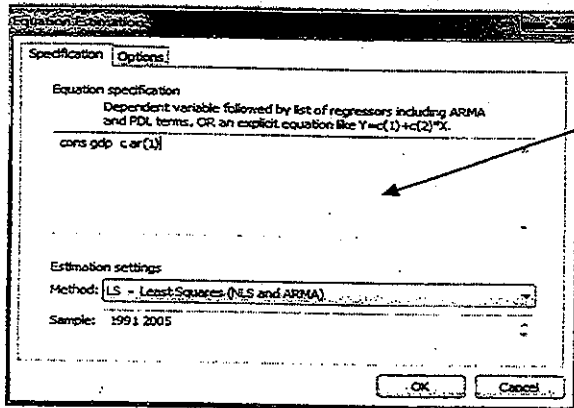
Kiểm định BG cho mô hình ta được kết quả :

#### Breusch-Godfrey Serial Correlation LM Test:

F-statistic	29.93650	Prob. F(1,12)	0.0001
Obs*R-squared	10.70780	Prob. Chi-Square(1)	0.0011

Vậy có thể thấy mô hình có tự tương quan bậc nhất, ta sẽ thử khắc phục bằng phương pháp Cochrane – Orcutt.

Để sử dụng Eviews khắc phục tự tương quan bậc nhất bằng Cochanre – Orcutt ta quay lại thủ tục ước lượng mô hình (hộp thoại Equation Specification). Ta khai báo thêm bậc của tự tương quan bằng cách gõ AR(j), j là bậc của tự tương quan, làm biến độc lập rồi nhấn OK.



Nếu là tự tương quan bậc 2, 3, 4 ... ta gõ vào AR(2), AR(3), AR(4) ...

Kết quả:

Dependent Variable: CONS  
 Method: Least Squares  
 Date: 05/10/14 Time: 00:41  
 Sample (adjusted): 1992 2005  
 Included observations: 14 after adjustments  
 Convergence achieved after 12 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP	0.146954	0.071722	2.048934	0.0651
C	590336.2	324962.7	1.816627	0.0966
AR(1)	0.950011	0.030173	31.48542	0.0000
R-squared	0.997872	Mean dependent var		279823.0
Adjusted R-squared	0.997485	S.D. dependent var		114130.9
S.E. of regression	5723.316	Akaike info criterion		20.32989
Sum squared resid	3.60E+08	Schwarz criterion		20.46683
Log likelihood	-139.3093	Hannan-Quinn criter.		20.31722
F-statistic	2579.283	Durbin-Watson stat		1.386023
Prob(F-statistic)	0.000000			
Inverted AR Roots	.95			

Kết quả ước lượng này được hiểu là kết quả ước lượng ở bước cuối cùng của phương pháp Cochanre - Orcutt. Hàm hồi quy mẫu thu được:

$$\text{CONS} = 0.146954 * \text{GDP} + 590336.2 + [\text{AR}(1) = 0.950011]$$

Với  $\text{AR}(1)=0.950011119033$  tức là  $e_t = 0.950011119033e_{t-1} + \varepsilon_t$ , nói cách khác ta thu được  $\hat{\rho} = 0.950011119033$  tức là tự tương quan gần như hoàn hảo.

### Khắc phục tự tương quan bằng cách đưa thêm biến

Hiện tượng tự tương quan còn là do bỏ sót biến gây ra. Lý thuyết về tiêu dùng và thu nhập cho rằng, tiêu dùng ở một thời điểm còn phụ thuộc vào chính bản thân nó ở thời kỳ trước, tức là trong mô hình nên chứa biến trễ của biến phụ thuộc. Ta ước lượng mô hình sau đây:

Dependent Variable: CONS  
 Method: Least Squares  
 Date: 05/10/14 Time: 00:51  
 Sample (adjusted): 1992 2005  
 Included observations: 14 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP	0.043099	0.033300	1.294249	0.2221
CONS(-1)	0.908655	0.063220	14.37292	0.0000
C	33186.73	4610.958	7.197361	0.0000
R-squared	0.997505	Mean dependent var		279823.0
Adjusted R-squared	0.997051	S.D. dependent var		114130.9
S.E. of regression	6197.711	Akaike info criterion		20.48916
Sum squared resid	4.23E+08	Schwarz criterion		20.62610
Log likelihood	-140.4241	Hannan-Quinn criter.		20.47648
F-statistic	2198.730	Durbin-Watson stat		1.335536
Prob(F-statistic)	0.000000			

Dùng tiêu chuẩn h - DW và BG để kiểm tra tự tương quan bậc nhất của mô hình trên ta được:

$$h = \left(1 - \frac{d}{2}\right) \sqrt{\frac{n}{1 - n \text{Var}(\hat{\alpha})}} = \left(1 - \frac{1,336}{2}\right) \sqrt{\frac{14}{1 - 14 \cdot 0,06322^2}} = 1,27851$$

Ta thấy:  $h < 1,96$  do đó chưa đủ cơ sở bác bỏ  $H_0$ , tức là mô hình này không có tự tương quan bậc nhất.

### Kết quả BG cũng tương tự:

#### Breusch-Godfrey Serial Correlation LM Test:

F-statistic	1.011483	Prob. F(1,10)	0.3383
Obs*R-squared	1.285999	Prob. Chi-Square(1)	0.2568

#### Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 05/10/14 Time: 00:57

Sample: 1992 2005

Included observations: 14

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP	0.005322	0.033701	0.157920	0.8777
CONS(-1)	-0.010808	0.064094	-0.168619	0.8695
C	569.1673	4643.171	0.122582	0.9049
RESID(-1)	0.307707	0.305955	1.005725	0.3383

R-squared	0.091857	Mean dependent var	2.24E-11
Adjusted R-squared	-0.180586	S.D. dependent var	5701.065
S.E. of regression	6194.478	Akaike info criterion	20.53566
Sum squared resid	3.84E+08	Schwarz criterion	20.71825
Log likelihood	-139.7496	Hannan-Quinn criter.	20.51876
F-statistic	0.337161	Durbin-Watson stat	1.814334
Prob(F-statistic)	0.798973		

#### Ước lượng sai số chuẩn theo Newey-West

Phương pháp này được đề xuất bởi Newey và West (1987), nó cho ta ước lượng sai số chuẩn của các hệ số hồi quy không dựa trên giả thiết về tự tương quan. Tuy các ước lượng không phải là ước lượng hiệu quả, nhưng sẽ là các ước lượng vững.

Tiếp tục dùng file ch6.2 là số liệu về tiêu dùng của người dân (CONS) và thu nhập (GDP) từ năm 1991 đến năm 2005 của Việt Nam. Kết quả hồi quy mô hình ban đầu:

Dependent Variable: CONS

Method: Least Squares

Date: 12/27/14 Time: 22:38

Sample: 1991 2005

Included observations: 15

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	64882.05	14347.45	4.522201	0.0006
GDP	0.527252	0.032631	16.15797	0.0000

R-squared	0.952569	Mean dependent var	265765.4
Adjusted R-squared	0.948920	S.D. dependent var	122717.9
S.E. of regression	27735.32	Akaike info criterion	23.42237
Sum squared resid	1.00E+10	Schwarz criterion	23.51677
Log likelihood	-173.6678	Hannan-Quinn criter.	23.42136
F-statistic	261.0800	Durbin-Watson stat	0.295616
Prob(F-statistic)	0.000000		

#### Ước lượng sai số chuẩn theo Newey-West:

[Equation] Estimate → Options → Heteroskedasticity... → Newey-West

Dependent Variable: CONS

Method: Least Squares

Date: 12/27/14 Time: 22:39

Sample: 1991 2005

Included observations: 15

Newey-West HAC Standard Errors & Covariance (lag truncation=2)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	64882.05	22113.58	2.934037	0.0116
GDP	0.527252	0.054593	9.657821	0.0000

R-squared	0.952569	Mean dependent var	265765.4
Adjusted R-squared	0.948920	S.D. dependent var	122717.9
S.E. of regression	27735.32	Akaike info criterion	23.42237
Sum squared resid	1.00E+10	Schwarz criterion	23.51677
Log likelihood	-173.6678	Hannan-Quinn criter.	23.42136
F-statistic	261.0800	Durbin-Watson stat	0.295616
Prob(F-statistic)	0.000000		

So sánh với bảng kết quả hồi quy ban đầu ta thấy các giá trị khác được giữ nguyên, chỉ có sai số chuẩn  $se(\beta_j)$  được điều chỉnh lớn lên so với các giá trị cũ.

### 6.3. Bài tập

**Bài 6.1.** Cho các số liệu về PCE – PCE services (chỉ tiêu dùng cá nhân cho dịch vụ) (tỷ USD) GDP – Gross domestic product (tổng sản phẩm quốc dân) (tỷ USD), FFR – Federal Funds Rate (lãi suất quỹ tiền tệ trung ương) (%), M2 – Money M2 (lượng cung tiền) (tỷ USD). Tiến hành khảo sát số liệu kinh tế Úc từ quý I năm 1983 đến quý I năm 2008.

#### Mô hình (1)

Dependent Variable: PCE

Method: Least Squares

Sample: 1983Q1 2008Q1

Included observations: 101

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP	0.407484	0.001967	207.1676	0.0000
C	-41.46932	16.89039	-2.455202	0.0158

R-squared	0.997699	Mean dependent var	3376.426
Adjusted R-squared	0.997675	S.D. dependent var	754.2616
S.E. of regression	36.36636	Akaike info criterion	10.04477
Sum squared resid	130928.7	Schwarz criterion	10.09655
Log likelihood	-505.2608	Hannan-Quinn criter.	10.06573
F-statistic	42918.42	Durbin-Watson stat	0.248554
Prob(F-statistic)	0.000000		

1. Dùng thống kê DW để kiểm tra hiện tượng tự tương quan của mô hình nói trên.
2. Cho biết đây là kiểm định gì? Thực hiện ra sao và kết quả thế nào.

#### Breusch-Godfrey Serial Correlation LM Test:

F-statistic	297.0427	Prob. F(1,98)	0.0000
Obs*R-squared	75.94448	Prob. Chi-Square(1)	0.0000

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 05/21/13 Time: 11:51

Sample: 1983Q1 2008Q1

Included observations: 101

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP	0.000280	0.000985	0.284396	0.7767
C	-2.068919	8.456272	-0.244661	0.8072
RESID(-1)	0.870810	0.050526	17.23493	0.0000

R-squared	0.751926	Mean dependent var	1.93E-13
Adjusted R-squared	0.746863	S.D. dependent var	36.18407
S.E. of regression	18.20520	Akaike info criterion	8.670544
Sum squared resid	32480.06	Schwarz criterion	8.748221
Log likelihood	-434.8625	Hannan-Quinn criter.	8.701990
F-statistic	148.5214	Durbin-Watson stat	1.911279
Prob(F-statistic)	0.000000		

#### Mô hình (2)

Dependent Variable: PCE

Method: Least Squares

Sample: 1983Q1 2008Q1

Included observations: 101

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP	0.370003	0.007799	47.44488	0.0000
FFR	-11.13448	1.651324	-6.742761	0.0000
M2	0.033806	0.009395	3.598469	0.0005
C	194.9818	35.11981	5.551904	0.0000

R-squared	0.998576	Mean dependent var	3376.426
Adjusted R-squared	0.998532	S.D. dependent var	754.2616
S.E. of regression	28.89912	Akaike info criterion	9.604298
Sum squared resid	81010.45	Schwarz criterion	9.707867
Log likelihood	-481.0170	Hannan-Quinn criter.	9.646226
F-statistic	22674.33	Durbin-Watson stat	0.359162
Prob(F-statistic)	0.000000		

3. Có tự tương quan trong mô hình nói trên hay không?

4. Dùng phương pháp BG để xem xét tự tương quan ta được kết quả sau. Cho biết kiểm định trên thực hiện thế nào. Cho kết quả ra sao?

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	188.4865	Prob. F(1,96)	0.0000
Obs*R-squared	66.91754	Prob. Chi-Square(1)	0.0000

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 05/21/13 Time: 11:53

Sample: 1983Q1 2008Q1

Included observations: 101

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP	0.001217	0.004555	0.267306	0.7898
FFR	1.832648	0.973441	1.882649	0.0628
M2	0.001008	0.005486	0.183721	0.8546
C	-24.16057	20.58261	-1.173834	0.2434
RESID(-1)	0.828160	0.060322	13.72904	0.0000

R-squared	0.662550	Mean dependent var	9.31E-14
Adjusted R-squared	0.648490	S.D. dependent var	28.46234
S.E. of regression	16.87484	Akaike info criterion	8.537762
Sum squared resid	27336.98	Schwarz criterion	8.667224
Log likelihood	-426.1570	Hannan-Quinn criter.	8.590172
F-statistic	47.12164	Durbin-Watson stat	1.965862
Prob(F-statistic)	0.000000		

Mô hình (3)

Dependent Variable: LOG(PCE)

Method: Least Squares

Date: 05/21/13 Time: 08:13

Sample: 1983Q1 2008Q1

Included observations: 101

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(GDP)	0.904916	0.030437	29.73062	0.0000
LOG(FFR)	-0.008427	0.002409	-3.498525	0.0007
LOG(M2)	0.060010	0.019509	3.076000	0.0027
C	-0.536344	0.120250	-4.460241	0.0000

R-squared	0.998075	Mean dependent var	8.099384
Adjusted R-squared	0.998015	S.D. dependent var	0.226979
S.E. of regression	0.010112	Akaike info criterion	-6.311301
Sum squared resid	0.009919	Schwarz criterion	-6.207732
Log likelihood	322.7207	Hannan-Quinn criter.	-6.269373
F-statistic	16760.95	Durbin-Watson stat	0.264734
Prob(F-statistic)	0.000000		

5. Việc đổi dạng mô hình có khắc phục được tự tương quan bậc nhất không?

**Bài 6.2.** File **ch6.2** là số liệu về tiêu dùng của người dân (CONS) và thu nhập (GDP) từ năm 1991 đến năm 2005 của Việt Nam.

1. Ước lượng mô hình:  $Cons = \beta_1 + \beta_2 GDP + u$  ?
2. Vẽ đồ thị Line Graph của RESID và đồ thị phân tán với trục hoành là thời gian và trục tung là RESID?
3. Vẽ đồ thị phân tán với trục hoành là RESID(-1) và trục tung là RESID và nêu nhận xét?
4. Tiến hành kiểm định Durbin-Watson d?

- Kiểm định Breusch – Godfrey về tự tương quan bậc nhất?
- Khắc phục tự tương quan dựa trên giá trị d của Durbin-Watson?
- Khắc phục tự tương quan dựa trên Durbin-Watson hai bước?
- Khắc phục tự tương quan bằng phương pháp Cochran – Orcutt ?

**Bài 6.3.** File **ch6.3.wfl** là số liệu về GDP (nghìn tỉ VNĐ), vốn (K: nghìn tỉ VNĐ) và lao động (L: nghìn người) của Bình Định trong giai đoạn 1995-2012 (theo giá so sánh 1994).

- Hãy ước lượng hàm sản xuất Cobb-Douglas:  $GDP = \beta_1 \cdot K^{\beta_2} \cdot L^{\beta_3} \cdot e^u$  và giải thích ý nghĩa các hệ số hồi quy ước lượng được?
- Bằng các phương pháp đã học hãy kiểm tra xem mô hình có tự tương quan bậc nhất hay không?
- Hãy đưa ra cách khắc phục các hiện tượng xảy ra ở câu 2?

**Bài 6.4.** Cho các số liệu đối với ngành dịch vụ của Mỹ từ năm 1948 – 1976 như sau:

Q: mức đầu; K – đầu vào vốn ; L – đầu vào lao động. tập số liệu **ch3.5.wfl**

- Hãy ước lượng các dạng hàm:

$$\ln Q = \beta_1 + \beta_2 \ln K + \beta_3 \ln L + u$$

$$Q = \beta_1 + \beta_2 \ln K + \beta_3 \ln L + u$$

$$\ln Q = \beta_1 + \beta_2 K + \beta_3 L + u$$

$$Q = \beta_1 + \beta_2 K + \beta_3 L + u$$

- Sau khi thu được các phần dư và phần dư chuẩn hoá (phần dư/ $\sigma$ ) từ mô hình hồi quy hãy vẽ đồ thị. Quan sát đồ thị có thể nói gì về tự tương quan bậc nhất của mô hình không?
- Bằng các phương pháp đã học hãy kiểm tra xem mô hình có tự tương quan bậc nhất trong các mô hình trên

hay không?

- Hãy đưa ra cách khắc phục các hiện tượng xảy ra ở câu 2?

Lấy bộ số liệu **ch5.8** với các biến số:

**loans** = commercial and industrial loans (billions of dollars, not seasonally adjusted) khoản vay cho công nghiệp và thương nghiệp (tỷ USD)

**deposits** = tổng mức tín dụng trong hệ thống ngân hàng (tỷ USD)

**Tbill3** = lợi suất trái phiếu kho bạc kỳ hạn 3 tháng (%)

**prime** = lãi suất cơ bản bình quân (%)

**aaa** = lãi lợi suất đáo hạn bình quân của trái phiếu dài hạn do bởi Moody's Investor Service (%)

**indprod** = index of industrial production, chỉ số sản xuất công nghiệp

Số liệu theo tháng từ tháng 01/1973 - 09/2003

Nguồn: Federal Reserve Board (<http://www.federalreserve.gov/releases/>)

- Hãy ước lượng mô hình sau

$$Loans = \beta_1 + \beta_2 Deposits + \beta_3 Tbill3 + \beta_4 Prime + \beta_5 Aaa + \beta_6 Indprod + u$$

- Dùng phương pháp quan sát đồ thị để kiểm tra tính tự tương quan của mô hình?
- Mô hình có tự tương quan bậc 1, 2, 3, 4, 5 hay không?
- Khắc phục tự tương quan bậc nhất bằng một trong số các phương pháp đã học
- Viết lại mô hình hồi quy với các hệ số ước lượng đã được khắc phục
- Với các hệ số ước lượng mới, hãy cho biết mức tăng tối đa của Loans khi deposits tăng 1?
- Khi cả hai Prime và Tbill3 giảm 1 thì Loans sẽ nằm trong khoảng nào?

## Chương 7: MỘT SỐ BÀI TẬP TỔNG HỢP

**Bài 7.1.** Bảng số liệu sau đây điều tra ở một số hộ gia đình.  $X_2$  là các khoản thu nhập từ lương và các khoản có tính chất lương;  $X_3$  là thu nhập ngoài lương;  $Y$  là chi tiêu. Đơn vị đo của các biến là triệu đồng.

Y	9	9	11	12	12.5	13	16	17	14	13.5
$X_2$	10	12	13	14	15	16	16	17	18	20
$X_3$	2	0	3	4	4	6	8	9	5	3

- Hãy ước lượng mô hình hồi quy  $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$  bằng phương pháp ma trận?
- Tính các giá trị thống kê khác, sau đó đối chiếu với kết quả tính toán bằng Eviews?
- Giải thích ý nghĩa kinh tế các hệ số hồi quy và hệ số xác định?
- Kiểm định sự phù hợp của hàm hồi quy?
- Các hệ số hồi quy của mô hình có ý nghĩa thống kê không?
- Khi thu nhập từ tiền lương tăng 1 triệu, thu nhập ngoài lương không đổi thì chi tiêu trung bình thay đổi trong khoảng nào?
- Có thể cho rằng cùng một mức tăng thu nhập như nhau, thu nhập ngoài lương tăng sẽ dẫn đến tăng chi tiêu cao hơn so với thu nhập từ lương tăng không?

- Với những người không có thu nhập khác ngoài lương, phải chăng khuynh hướng tiêu dùng chưa đến 0,8?
- Nếu cả thu nhập từ lương và thu nhập ngoài lương cùng tăng 1 triệu thì chi tiêu trung bình thay đổi thế nào? Tiết kiệm thay đổi thế nào?
- Dự báo mức chi tiêu trung bình của hộ có thu nhập từ lương là 20 triệu đồng, thu nhập ngoài lương là 10 triệu đồng?

**Bài 7.2.** Cho mô hình hồi quy  $Y_i = \beta_1 + \beta_2 X_i + u_i$ . Hãy cho biết:

- Nếu nhân 10 với mỗi  $X_i$  thì khi đó các  $e_i$  và các giá trị ước lượng  $\hat{Y}_i$  có thay đổi hay không? Vì sao?
- Nếu ta cộng vào mỗi  $X_i$  với một hằng số (chẳng hạn 100) thì khi đó các  $e_i$  và các giá trị ước lượng  $\hat{Y}_i$  có thay đổi hay không? Vì sao?

**Bài 7.3.** Cho các biến số với bộ số liệu được lấy theo tháng trong một doanh nghiệp:

Q: Sản lượng

TR: Tổng doanh thu

T: Lượng thuế phải đóng

W: Lương của lao động.

- Có ý kiến cho rằng lương của lao động phụ thuộc vào sản lượng của doanh nghiệp, tổng doanh thu, và lượng thuế phải đóng; khi đó lương của lao động sẽ tăng 2 đơn vị khi sản lượng tăng 2 đơn vị, nhưng sẽ giảm 10 đơn vị khi lượng thuế phải đóng tăng 1%. Hãy xây dựng mô hình kinh tế lượng và trình bày cách phân tích các nhận định trên?
- Khi cả doanh thu và lượng thuế phải đóng cùng tăng một lượng thì dường như yếu tố doanh thu tác động tới lương của lao động mạnh hơn so với thuế, khi đó lương của lao động là tăng. Hãy nêu cách phân tích

nhận định trên (chi tiết công thức tính toán)?

**Bài 7.4.** Người ta muốn khảo sát lượng cầu (VND) về Trà xanh Không độ và Trà xanh 100 của người dân thành phố Quy Nhơn theo giá bán (VND) của chúng. Biết là khi giá bán tăng 1 đơn vị thì lượng cầu của cả hai loại đều giảm 10%.

1. Hãy xây dựng mô hình và nêu cách kiểm tra nhận định trên
2. Có người cho rằng với cùng một mức giá thì lượng cầu trung bình về hai loại này chênh nhau khoảng 5%. Hãy nêu cách kiểm tra nhận định này.

**Bài 7.5.** Người ta muốn khảo sát lượng cầu (VND) về Cafe và Trà của người dân thành phố Quy Nhơn theo giá bán (VND) của chúng. Biết là khi giá bán tăng 1% đơn vị thì lượng cầu của cả hai loại đều giảm 9%.

1. Hãy xây dựng mô hình và nêu cách kiểm tra nhận định trên
2. Có người cho rằng với cùng một mức giá thì lượng cầu trung bình về hai loại này chênh nhau khoảng 8%. Hãy nêu cách kiểm tra nhận định này.

**Bài 7.6.** Cho các biến số: Q – lượng cầu thịt lợn (10kg/người); Y – thu nhập/đầu người (USD); P – giá một kg thịt lợn (USD). Kết quả Eviews như sau:

Dependent Variable: Q  
Included observations: 30

Variable	Coefficient	Std. Error	t-Statistic	Prob.
P		0.259020	-2.968379	0.0062
Y	0.038991	0.005757	6.773213	0.0000
C	1.368933	3.298834	0.414975	0.6814
R-squared	0.671073	Mean dependent var		10.89448

Cho  $cov(\hat{\beta}_P, \hat{\beta}_Y) = 1,45E - 05$ .

1. Khi giá tăng 1USD thì lượng cầu thịt lợn giảm tối đa bao nhiêu?
2. Có phải giá bán tác động lên lượng cầu thịt lợn yếu hơn so với yếu tố thu nhập?

**Bài 7.7.** File **ch7.7** cho số liệu về doanh thu (Y), chi phí quảng cáo (X2) và tiền lương của nhân viên tiếp thị (X3) của 12 công ty tư nhân (đơn vị tính của các biến là triệu đ/tháng).

Y	X2	X3
126	17	11
148	23	14
105	18	9
162	22	16
101	14	9
175	24	17
160	23	15
127	15	11
138	16	12
143	21	14
158	22	15
137	13	13



1. Ước lượng hàm hồi quy dạng log-log và viết hàm hồi quy gốc. Nêu ý nghĩa kinh tế các hệ số hồi quy?
2. Khi vốn tăng thêm 1%, giữ nguyên lượng lao động, sản lượng tối đa có thể tăng thêm bao nhiêu với độ tin cậy 95%?
3. Khi lượng lao động tăng thêm 1%, giữ nguyên lượng vốn, sản lượng có thể tăng thêm ít nhất là bao nhiêu % với độ tin cậy 95%?
4. Với mức ý nghĩa 5%, có thể cho rằng với cùng một tỷ lệ, tăng quy mô vốn đem lại hiệu quả cũng bằng như tăng quy mô về lao động không?
5. Theo bạn, việc tăng quy mô sản xuất của \$30\$ doanh nghiệp trên có thể làm tăng hiệu quả không với độ tin cậy 95%?

**Bài 7.8.** File ch7.8 là số liệu về tăng trưởng GDP (GGDP), tăng trưởng nông nghiệp (GAGR), tăng trưởng công nghiệp (GIND) và tăng trưởng dịch vụ (GSER) của Việt Nam các năm trong giai đoạn 1990-2012 (đơn vị các biến: %).

obs	GAGR	GGDP	GIND	GSER
1990	1	5.1	2.3	10.2
1991	2.2	5.8	7.7	7.4
1992	6.9	8.7	12.8	7.6
1993	3.3	8.1	12.6	8.6
1994	3.4	8.8	13.4	9.6
1995	4.8	9.5	13.6	9.8
1996	4.4	9.3	14.5	8.8
1997	4.3	8.2	12.6	7.1
1998	3.5	5.8	8.3	5.1
1999	5.2	4.8	7.7	2.3
2000	4.6	6.8	10.1	5.3

2001	3	6.9	10.4	6.1
2002	4.2	7.1	9.5	6.5
2003	3.6	7.3	10.5	6.5
2004	4.4	7.8	10.2	7.3
2005	4	8.4	10.7	8.5
2006	3.7	8.2	10.4	8.3
2007	3.8	8.5	10.2	8.9
2008	4.7	6.3	6	7.4
2009	1.8	5.3	5.5	6.6
2010	2.8	6.8	7.7	7.5
2011	4	5.9	5.5	7
2012	2.7	5	4.5	6.4

1. Vẽ đồ thị Line Graph của các biến trên cùng hệ trục tọa độ?
2. Vẽ đồ thị phân tán của biến GGDP theo từng biến và cho nhận xét?
3. Ước lượng mô hình hồi quy tuyến tính thông thường của GGDP theo các biến còn lại và giải thích ý nghĩa các hệ số hồi quy?
4. Mô hình có được coi là phù hợp không?
5. Trong điều kiện các yếu tố khác không đổi, khi nông nghiệp tăng trưởng 1% thì GDP tăng trưởng tối đa là bao nhiêu?
6. Khi các ngành kinh tế cùng tăng trưởng 1% thì GDP tăng trưởng trong khoảng nào?
7. Phải chăng tăng trưởng dịch vụ ảnh hưởng đến tăng trưởng GDP mạnh hơn là tăng trưởng nông nghiệp?
8. Đặt  $t$  là biến xu thế ( $t = 1$  năm 1990, ...,  $t = 23$  năm 2012). Có nên đưa thêm biến  $t$  vào mô hình ở câu 3 hay không? Vì sao?

**Bài 7.9.** File **ch7.9** là số liệu về tiết kiệm và tiêu dùng của một quốc gia trong giai đoạn 1970-1995, trong đó:

SAVING: Tiết kiệm quốc gia tính bằng tỉ USD

INCOME: Thu nhập quốc gia tính bằng tỉ USD

- Giai đoạn 1970-1981 chính sách tiền tệ thắt chặt và mức lãi suất rất cao, còn giai đoạn 1982-1995 chính sách tiền tệ nới lỏng và lãi suất thấp. Hãy ước lượng hàm hồi tuyến tính của SAVING theo INCOME (có hệ số chặn) cho từng giai đoạn và giải thích ý nghĩa của hệ số góc?
- Hãy ước lượng hàm hồi quy dạng bội có dạng sau đây cho cả giai đoạn 1970-1995:

$$Saving_t = \beta_1 + \beta_2 Dum_t + \beta_3 Income_t + u_t$$

Với DUM = 0 trong giai đoạn 1970-1981,

DUM = 1 trong giai đoạn 1982-1995.

Giải thích ý nghĩa các hệ số hồi quy riêng trong hàm hồi quy?

- Từ câu 2 anh chị hãy viết phương trình hồi quy cho trường hợp DUM = 1 và DUM = 0. So sánh kết quả này với kết quả mà anh chị đã tìm ra ở câu 1. Nêu nhận xét của anh chị về kết quả tìm được?
- Hãy ước lượng hàm hồi quy có dạng sau đây cho giai đoạn 1970-1995:

$$Saving_t = \beta_1 + \beta_2 Dum_t + \beta_3 Income_t + \beta_4 Dum * Income_t + u_t$$

- Từ kết quả của câu 4, hãy cho biết giữa 2 giai đoạn có sự khác biệt có ý nghĩa thống kê của hệ số hồi quy hay không? Sự khác biệt này được giải thích bằng ý nghĩa kinh tế là 2 giai đoạn chính sách khác nhau ảnh hưởng thế nào đến hành vi tiết kiệm của công chúng?

**Bài 7.10.** File **ch7.10** là số liệu về các biến như sau:

BEER: mức chi tiêu cho bia/người/năm

Y: thu nhập/người/năm

AGE: tuổi

SEX: giới tính (1: nam, 0: nữ)

- D1 = 1: nếu tốt nghiệp trung học, D1 = 0 đối với các trường hợp khác
- D2 = 1: nếu tốt nghiệp cao đẳng, D2 = 0 đối với các trường hợp khác
- D3 = 1: nếu tốt nghiệp đại học, D3 = 0 đối với các trường hợp khác

- Hãy ước lượng hàm tiêu dùng về bia phụ thuộc vào thu nhập và tuổi?
- Nhu cầu về bia có phụ thuộc giới tính không? Vì sao?
- Trình độ học vấn có ảnh hưởng đến nhu cầu về bia hay không? Vì sao?
- Hãy ước lượng nhu cầu về bi phụ thuộc vào thu nhập, giới tính, tuổi và trình độ học vấn?
- Tìm ma trận tương quan của các biến BEER, Y, SEX, AGE. Biến nào có ảnh hưởng nhiều nhất đến nhu cầu về bia? Vì sao?
- Hãy ước lượng nhu cầu về bia theo các biến AGE, Y, SEX và dùng mô hình này để dự báo nhu cầu về bia trung bình của một người có giới tính là nam, 50 tuổi và có thu nhập là 30000?
- Mô hình ước lượng ở câu 6 có xảy ra hiện tượng cộng tuyến, phương sai thay đổi, tự tương quan không?
- Vẽ đồ thị Line Graph của BEER và BEERF trên cùng hệ trục tọa độ và nhận xét?

**Bài 7.11.** File **ch7.11** chứa số liệu của 51 quan sát của các biến được giải thích như sau:

MORT:	Tỉ lệ tử vong chung trên 100.000 dân số
INCC:	Thu nhập đầu người tính bằng USD
POV:	Tỉ lệ người dân trong nước sống dưới mức nghèo khó
EDU1:	Tỉ lệ dân số đã học trung học
EDU2:	Tỉ lệ dân số đã học trung học và đại học
ALCC:	Tiêu dùng cồn (rượu) tính bằng lít trên đầu người
TOBC:	Tiêu dùng thuốc lá đầu người tính bằng gói
HEXC:	Chỉ tiêu y tế bình quân đầu người (USD)
URB:	Tỉ lệ dân số sống tại các khu vực thành thị
AGED:	Tỉ lệ dân số có độ tuổi trên 65
PHYS:	Các cán bộ y tế trên 1.000.000 dân

1. Chọn biến Mort là biến phụ thuộc. Đối với từng biến giải thích, hãy lý giải tại sao nó có thể có tác động lên tỉ lệ tử vong tổng thể và chỉ ra hướng tác động này?
2. Ước lượng mô hình hồi quy đa biến bao gồm tất cả những biến giải thích đang có và thực hiện kiểm định sự phù hợp của mô hình ở mức ý nghĩa 5%?
3. Với mức ý nghĩa 10%, hãy xác định các biến mà anh chị có thể muốn bỏ qua?
4. Đầu tiên hãy thực hiện một kiểm định WALD đối với việc bỏ tất cả những biến này. Sau đó, bỏ tất cả chúng và ước lượng lại mô hình. Hãy đánh giá kết quả của mô hình hồi quy mới này?
5. Trở lại với mô hình gốc ban đầu, lần lượt mỗi lần bỏ

một biến không có ít ý nghĩa thống kê nhất và ước lượng lại mô hình. Hãy cho nhận xét về mô hình hồi quy mới?

6. Kết thúc quá trình thực hiện ở câu 5 ta thu được mô hình cuối cùng, mô hình này có giống mô hình thu được từ kết quả câu 4 hay không? Từ kết quả đó, bạn có thể đưa ra kết luận gì?

**Bài 7.12.** File **ch7.12** là số liệu hàng năm về sản xuất cá ngừ trắng tại vùng Basque Tây Ban Nha trong giai đoạn 1961-1994. Trong đó:

CATCH: Tổng lượng cá đánh bắt tính bằng đơn vị 1000 tấn  
EFFORT: Tổng số ngày đánh cá theo đơn vị nghìn ngày.

1. Hãy ước lượng các mô hình hồi quy sau:

$$Catch_i = \beta_1 + \beta_2 Effort_i + u_i \quad (1)$$

$$Catch_i = \beta_1 + \beta_2 Effort_i + \beta_3 Effort_i^2 + u_i \quad (2)$$

2. Giải thích các hệ số  $\beta_2, \beta_3$  ước lượng được trong câu trên?
3. Nếu mô hình đúng là mô hình (2), nhưng thay vì chọn mô hình đó anh chị lại chọn mô hình (1) để thực hiện phân tích về sản lượng sản xuất cá ngừ trắng, thì lúc này ta gặp phải những khó khăn gì? Hãy giải thích và thực hiện những kiểm định cần thiết?

**Bài 7.13.** File **ch7.13** cho số liệu từ tháng 4/2001 đến tháng 12/2003 về các biến như sau:

Y: Lượng tiền huy động (triệu VNĐ), X2: Khoản tín dụng (triệu VNĐ), X3: Chi phí quảng cáo (triệu VNĐ), X4: Tỷ giá VNĐ/USD, X5: Giá vàng 9999 (triệu VNĐ/lượng).

1. Tính hệ số tương quan của Y với tất cả các biến còn lại.

Tìm hàm hồi quy tuyến tính mẫu của Y với biến nào có hệ số trong quan tuyến tính cao nhất. Kiểm định xem mô hình này có bị bỏ sót biến cần thiết hay không?

- Hồi quy Y theo các biến X2, X3, X4, X5 và kiểm định xem mô hình có bị thừa biến hay không?
- Định mẫu từ tháng 4/2001 đến tháng 8/2003. Sau đó tìm mô hình hồi quy tuyến tính mẫu dùng để dự báo tốt nhất?

**Bài 7.14.** File **ch7.14** cho số liệu về sản lượng điện tiêu thụ từng tháng tại Bình Định (Y) trong giai đoạn từ tháng 1/2008 đến tháng 12/2013 (đơn vị: KW).

- Dùng mẫu từ tháng 1/2008 đến tháng 12/2012 để hồi quy Y theo biến xu thế T1 ( $T1 = 1$ ) ứng với tháng 1/2008, ...,  $T1 = 60$  ứng với tháng 12/2012). Sử dụng mô hình hồi quy đã ước lượng để dự báo lượng sản điện tiêu thụ của 12 tháng năm 2013. Từ đó tính các chỉ tiêu đánh giá độ chính xác của mô hình?
- Sử dụng mẫu từ quý I/2008 đến quý IV/2012 để hồi quy Y2 theo T2 và D2, D3, D4, trong đó Y2 là sản lượng điện tiêu thụ theo từng quý; T2 là biến xu thế ( $T2 = 1$  ứng với quý I/2008, ...,  $T2 = 20$  ứng với quý IV/2012; D2, D3, D4 là các biến giả:

D2 = 1 nếu quan sát ở quý II, D2 = 0 nếu quan sát ở quý khác

D3 = 1 nếu quan sát ở quý III, D3 = 0 nếu quan sát ở quý khác

D4 = 1 nếu quan sát ở quý IV, D2 = 0 nếu quan sát ở quý khác.

- Dùng mô hình đã ước lượng được ở câu 2 để dự báo sản lượng điện tiêu thụ của quý IV/2013 và tính các chỉ tiêu đánh giá độ chính xác của mô hình?
- Nên dùng mô hình nào trong hai mô hình ở câu 1 và câu 2 để dự báo sản lượng điện tiêu thụ? Vì sao?

**Bài 7.15.** File **ch7.15** là số liệu về doanh số và tồn kho của công nghiệp một quốc gia Châu Âu trong giai đoạn 1950-1991.

SALES: Doanh số của ngành công nghiệp (triệu \$)

INVENTORIES: Giá trị tồn kho của ngành công nghiệp (triệu \$)

- Hồi quy mô hình tuyến tính có hệ số chặn của SALES theo INVENTORIES. Sau đó, vẽ đồ thị phần dư của mô hình hồi quy này theo INVENTORIES. Dựa vào đồ thị đó, anh chị có ý kiến gì về vấn đề trong quan chuỗi?
- Hãy trình bày kiểm định Durbin – Watson xem phần dư của phương trình trên có tương quan chuỗi bậc nhất hay không?
- Đề nghị giải pháp khắc phục tương quan chuỗi trong phép hồi quy nói trên nếu có?

**Bài 7.16.** File **ch7.16** là số liệu về điểm học phần Kinh tế lượng (DIEM), điểm thi đầu vào đại học môn Toán (DIEMDH), thời gian tự học môn Kinh tế lượng (TUHOC: giờ/tuần) và khối thi đầu vào (KHOI) của 200 sinh viên khối ngành Kinh tế.

- Hồi quy mô hình  $Diem_i = \beta_1 + \beta_2 Tuhoc_i + \beta_3 DiemDh_i + u_i$  (1).
- Phải chăng cả thời gian tự học và điểm thi vào đại học môn Toán đều không ảnh hưởng đến điểm thi môn Kinh tế lượng của sinh viên?
- Nếu sinh viên có thêm 2 giờ tự học/tuần thì kết quả trung bình môn Kinh tế lượng sẽ tăng lên ít nhất bao nhiêu điểm?
- Ngoài các yếu tố về thời gian tự học, điểm thi đầu vào đại học thì khối thi đại học (khối A hoặc khối D) có ảnh hưởng đến kết quả môn Kinh tế lượng của sinh

viên không? Nếu có thì khối nào đạt kết quả cao hơn?

- Cùng một điểm thi đại học môn Toán như nhau, hiệu quả 1 giờ tự học của sinh viên thi vào khối A cao hơn sinh viên thi vào khối D. Hãy kiểm định nhận định trên?

**Bài 7.17.** Cho các biến số: GDP – tổng sản phẩm quốc nội; I – đầu tư; CPI – chỉ số giá tiêu dùng; R – tỷ lệ lạm phát.  
Hồi quy thu được kết quả sau đây.

Dependent Variable: I  
Method: Least Squares  
Sample: 1968 1982  
Included observations: 15

Variable	Coefficient	Std. Error	t-Statistic
GDP	0.689021	0.064034	10.76025
CPI		1.137803	-8.391809
R		2.296132	-1.834127
C	357.1887	42.73375	8.358469
R-squared	0.992757	Mean dependent var	276.0067
Sum squared resid	1401.881		
F-statistic	502.5950	Durbin-Watson stat	1.917191
Prob(F-statistic)	0.000000		

Cho biết  $COV(\hat{\beta}_{CPI}, \hat{\beta}_R) = 0,509$ ;  $COV(\hat{\beta}_{CPI}, \hat{\beta}_{GDP}) = -0,072$ .

- Cho biết mối quan hệ trong thực tế giữa từng biến độc lập với biến phụ thuộc.
- Chỉ số giá tiêu dùng tăng 1 đơn vị thì đầu tư giảm ít hơn 8 đơn vị phải không?
- Có phải CPI tác động tới đầu tư mạnh hơn so với tỷ lệ lạm phát?
- Khi cả GDP và CPI cùng giảm 1 thì I sẽ tăng hay giảm trong khoảng nào?

**Bài 7.18.** Cho các biến số: Y – sản lượng; L lao động; K – vốn.  
Khảo sát 30 doanh nghiệp.

Kết quả hồi quy như sau:

Dependent Variable: LOG(Y/L)  
Included observations: 30

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(K/L)	0.735804	0.065794	11.18348	0.0000
LOG(L)	0.684798	0.098532	6.950017	0.0000
C	0.424816	0.137808	3.082671	0.0047
R-squared	0.831003	Mean dependent var		-0.286039
Adjusted R-squared	0.818485	S.D. dependent var		1.010559
F-statistic	66.38305	Durbin-Watson stat		2.192857
Prob(F-statistic)	0.000000			

và ma trận hiệp phương sai:

	LOG(K/L)	LOG(L)	C
LOG(K/L)	0.004329	0.005040	0.005443
LOG(L)	0.005040	0.009709	0.011125
C	0.005443	0.011125	0.018991

- Cho biết khi lao động giảm đi 1,5 lần so với trước thì năng suất lao động trung bình giảm đi 2,25 lần so với trước phải không?
- Khi K/L tăng 1% thì năng suất lao động cũng tăng 1% phải không?

**Bài 7.19.** Cho các biến số sau đây:

Y : chiều cao của sinh viên (cm). X2 : chỉ tiêu cho ăn uống (đồng). X3 chiều cao của bố (cm), Chế độ tập thể dục (đều đặn, thỉnh thoảng, hiếm khi).

D1 = 1 nếu tập đều đặn  
= 0 nếu chế độ khác

D2 = 1 nếu hiếm khi tập  
= 0 nếu chế độ khác

Mô hình (1)

(1) Dependent Variable: Y  
Included observations: 70

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X2	1.93E-06	1.76E-06	1.095624	0.2773
X3	0.217496	0.135122	1.609622	0.1123
D1	9.108427	2.338937	3.894259	0.0002
D2	-3.611370	1.982937	-1.821222	0.0732
C	121.9822	22.19308	5.496404	0.0000
R-squared	0.273504	Mean dependent var		159.8429
Adjusted R-squared	0.228797	S.D. dependent var		6.956669
Log likelihood	-223.4178	F-statistic		6.117653
Durbin-Watson stat	2.001041	Prob(F-statistic)		0.000306

1. Có phải X2 tác động tới Y yếu hơn so với X3 không?

$$\text{Cho } \text{cov}(\hat{\beta}_{X2}, \hat{\beta}_{X3}) = 6,42E - 09.$$

2. Khi chi tiêu cho ăn uống tăng thì chiều cao của sinh viên có tăng không? Tăng tối đa là bao nhiêu?

3. Với các biến số Y và X3 như mô hình (1). Ta ước lượng mô hình (2):

(2) Dependent Variable: LOG(Y)  
Included observations: 70

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X3	0.001875	0.000724	2.591540	0.0110
C	4.762632	0.119410	39.88466	0.0000
R-squared	0.064136	Mean dependent var		5.071904
Adjusted R-squared	0.054586	S.D. dependent var		0.042245
S.E. of regression	0.041075	Akaike info criterion		-3.527019
Sum squared resid	0.165344	Schwarz criterion		-3.474915
Log likelihood	178.3509	F-statistic		6.716080
Durbin-Watson stat	1.789888	Prob(F-statistic)		0.011014

- Có thể nói gì về ý nghĩa của hệ số góc gắn với mô hình (2)?
- Mô hình này chỉ có mức độ phù hợp là khoảng 6,4% (rất thấp) nhưng kiểm định phù hợp lại thấy mô hình hồi quy là phù hợp. Có gì mâu thuẫn ở đây không?
- Khi X3 tăng 1 đơn vị thì chiều cao tăng tối đa là bao nhiêu?

Bài 7.20. File ch7.20 là số liệu về giá trị nhập khẩu (IMPORTS: triệu USD), GDP (tỉ VNĐ) và chỉ số giá tiêu dùng (CPI, 1994 = 100) trong thời kỳ 1995-2013.

- Hãy ước lượng mô hình tuyến tính log của IMPORTS theo các biến còn lại?
- Từ kết quả ước lượng trên, bạn có nghi ngờ gì về hiện tượng đa cộng tuyến không? Vì sao?
- Hãy hồi quy các mô hình sau:

$$\ln(\text{IMPORT}) = \alpha_1 + \alpha_2 \ln(\text{GDP}) + u \quad (1)$$

$$\ln(\text{IMPORT}) = \lambda_1 + \lambda_2 \ln(\text{CPI}) + u \quad (2)$$

$$\ln(\text{GDP}) = \gamma_1 + \gamma_2 \ln(\text{CPI}) + u \quad (3)$$

Từ các mô hình này, bạn có nhận xét gì về hiện tượng đa cộng tuyến trong số liệu. Từ đó đưa ra cách khắc phục và chọn mô hình tốt nhất?

- Từ mô hình lựa chọn trong câu 3, bạn hãy kiểm tra xem có hiện tượng tự tương quan xảy ra không? Nếu có hãy chỉ ra cách khắc phục?
- Kiểm tra tính chuẩn của nhiễu trong mô hình lựa chọn ở câu 3?

## PHỤ LỤC CÁC BẢNG THỐNG KÊ

**Bảng 1. Giá trị hàm mật độ phân phối chuẩn hóa**

$$\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{(-u^2/2)}$$

u	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.3989	0.3950	0.3910	0.3872	0.3833	0.3795	0.3757	0.3720	0.3683	0.3646
0.1	0.3610	0.3574	0.3538	0.3503	0.3468	0.3434	0.3400	0.3366	0.3332	0.3299
0.2	0.3266	0.3234	0.3202	0.3170	0.3138	0.3107	0.3076	0.3045	0.3015	0.2985
0.3	0.2955	0.2926	0.2897	0.2868	0.2840	0.2811	0.2783	0.2756	0.2728	0.2701
0.4	0.2674	0.2648	0.2621	0.2595	0.2569	0.2544	0.2518	0.2493	0.2469	0.2444
0.5	0.2420	0.2396	0.2372	0.2348	0.2325	0.2302	0.2279	0.2256	0.2234	0.2211
0.6	0.2189	0.2168	0.2146	0.2125	0.2104	0.2083	0.2062	0.2041	0.2021	0.2001
0.7	0.1981	0.1961	0.1942	0.1923	0.1903	0.1884	0.1866	0.1847	0.1829	0.1811
0.8	0.1793	0.1775	0.1757	0.1740	0.1722	0.1705	0.1688	0.1671	0.1655	0.1638
0.9	0.1622	0.1606	0.1590	0.1574	0.1558	0.1543	0.1528	0.1512	0.1497	0.1482
1.0	0.1468	0.1453	0.1439	0.1424	0.1410	0.1396	0.1382	0.1368	0.1355	0.1341
1.1	0.1328	0.1315	0.1302	0.1289	0.1276	0.1263	0.1251	0.1238	0.1226	0.1214
1.2	0.1202	0.1190	0.1178	0.1166	0.1154	0.1143	0.1132	0.1120	0.1109	0.1098
1.3	0.1087	0.1076	0.1066	0.1055	0.1045	0.1034	0.1024	0.1014	0.1004	0.0994
1.4	0.0984	0.0974	0.0964	0.0955	0.0945	0.0936	0.0926	0.0917	0.0908	0.0899
1.5	0.0890	0.0881	0.0873	0.0864	0.0855	0.0847	0.0838	0.0830	0.0822	0.0814
1.6	0.0805	0.0797	0.0790	0.0782	0.0774	0.0766	0.0759	0.0751	0.0744	0.0736
1.7	0.0729	0.0722	0.0714	0.0707	0.0700	0.0693	0.0686	0.0680	0.0673	0.0666
1.8	0.0659	0.0653	0.0646	0.0640	0.0634	0.0627	0.0621	0.0615	0.0609	0.0603
1.9	0.0597	0.0591	0.0585	0.0579	0.0573	0.0568	0.0562	0.0556	0.0551	0.0545
2.0	0.0540	0.0535	0.0529	0.0524	0.0519	0.0514	0.0508	0.0503	0.0498	0.0493
2.1	0.0489	0.0484	0.0479	0.0474	0.0469	0.0465	0.0460	0.0456	0.0451	0.0446
2.2	0.0442	0.0438	0.0433	0.0429	0.0425	0.0420	0.0416	0.0412	0.0408	0.0404
2.3	0.0400	0.0396	0.0392	0.0388	0.0384	0.0380	0.0377	0.0373	0.0369	0.0366
2.4	0.0362	0.0358	0.0355	0.0351	0.0348	0.0344	0.0341	0.0337	0.0334	0.0331

2.5	0.0327	0.0324	0.0321	0.0318	0.0315	0.0312	0.0308	0.0305	0.0302	0.0299
2.6	0.0296	0.0293	0.0290	0.0288	0.0285	0.0282	0.0279	0.0276	0.0274	0.0271
2.7	0.0268	0.0265	0.0263	0.0260	0.0258	0.0255	0.0252	0.0250	0.0247	0.0245
2.8	0.0243	0.0240	0.0238	0.0235	0.0233	0.0231	0.0228	0.0226	0.0224	0.0222
2.9	0.0220	0.0217	0.0215	0.0213	0.0211	0.0209	0.0207	0.0205	0.0203	0.0201
3.0	0.0199	0.0197	0.0195	0.0193	0.0191	0.0189	0.0187	0.0185	0.0183	0.0182
3.1	0.0180	0.0178	0.0176	0.0174	0.0173	0.0171	0.0169	0.0168	0.0166	0.0164
3.2	0.0163	0.0161	0.0159	0.0158	0.0156	0.0155	0.0153	0.0152	0.0150	0.0149
3.3	0.0147	0.0146	0.0144	0.0143	0.0141	0.0140	0.0139	0.0137	0.0136	0.0134
3.4	0.0133	0.0132	0.0131	0.0129	0.0128	0.0127	0.0125	0.0124	0.0123	0.0122
3.5	0.0120	0.0119	0.0118	0.0117	0.0116	0.0115	0.0113	0.0112	0.0111	0.0110
3.6	0.0109	0.0108	0.0107	0.0106	0.0105	0.0104	0.0103	0.0102	0.0101	0.0100
3.7	0.0099	0.0098	0.0097	0.0096	0.0095	0.0094	0.0093	0.0092	0.0091	0.0090
3.8	0.0089	0.0088	0.0087	0.0087	0.0086	0.0085	0.0084	0.0083	0.0082	0.0082
3.9	0.0081	0.0080	0.0079	0.0078	0.0078	0.0077	0.0076	0.0075	0.0075	0.0074
4.0	0.0073	0.0072	0.0072	0.0071	0.0070	0.0070	0.0069	0.0068	0.0067	0.0067
4.1	0.0066	0.0065	0.0065	0.0064	0.0064	0.0063	0.0062	0.0062	0.0061	0.0060
4.2	0.0060	0.0059	0.0059	0.0058	0.0057	0.0057	0.0056	0.0056	0.0055	0.0055
4.3	0.0054	0.0054	0.0053	0.0053	0.0052	0.0051	0.0051	0.0050	0.0050	0.0049
4.4	0.0049	0.0048	0.0048	0.0048	0.0047	0.0047	0.0046	0.0046	0.0045	0.0045

Bảng 1. Giá trị hàm mật độ phân phối chuẩn hóa

$$\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{(-u^2/2)} \quad (tt)$$

u	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
4.5	0.0044	0.0044	0.0043	0.0043	0.0043	0.0042	0.0042	0.0041	0.0041	0.0041
4.6	0.0040	0.0040	0.0039	0.0039	0.0039	0.0038	0.0038	0.0037	0.0037	0.0037
4.7	0.0036	0.0036	0.0036	0.0035	0.0035	0.0035	0.0034	0.0034	0.0033	0.0033
4.8	0.0033	0.0033	0.0032	0.0032	0.0032	0.0031	0.0031	0.0031	0.0030	0.0030
4.9	0.0030	0.0029	0.0029	0.0029	0.0029	0.0028	0.0028	0.0028	0.0027	0.0027
5.0	0.0027	0.0027	0.0026	0.0026	0.0026	0.0026	0.0025	0.0025	0.0025	0.0025
5.1	0.0024	0.0024	0.0024	0.0024	0.0023	0.0023	0.0023	0.0023	0.0022	0.0022
5.2	0.0022	0.0022	0.0022	0.0021	0.0021	0.0021	0.0021	0.0021	0.0020	0.0020
5.3	0.0020	0.0020	0.0020	0.0019	0.0019	0.0019	0.0019	0.0019	0.0018	0.0018
5.4	0.0018	0.0018	0.0018	0.0017	0.0017	0.0017	0.0017	0.0017	0.0017	0.0016
5.5	0.0016	0.0016	0.0016	0.0016	0.0016	0.0016	0.0015	0.0015	0.0015	0.0015
5.6	0.0015	0.0015	0.0014	0.0014	0.0014	0.0014	0.0014	0.0014	0.0014	0.0013
5.7	0.0013	0.0013	0.0013	0.0013	0.0013	0.0013	0.0013	0.0012	0.0012	0.0012
5.8	0.0012	0.0012	0.0012	0.0012	0.0012	0.0011	0.0011	0.0011	0.0011	0.0011
5.9	0.0011	0.0011	0.0011	0.0011	0.0011	0.0010	0.0010	0.0010	0.0010	0.0010
6.0	0.0010	0.0010	0.0010	0.0010	0.0010	0.0009	0.0009	0.0009	0.0009	0.0009
6.1	0.0009	0.0009	0.0009	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008
6.2	0.0008	0.0008	0.0008	0.0008	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
6.3	0.0007	0.0007	0.0007	0.0007	0.0007	0.0007	0.0007	0.0007	0.0007	0.0007
6.4	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006
6.5	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005



6.6	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005
6.7	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0004
6.8	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
6.9	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
7.0	0.0004	0.0004	0.0004	0.0004	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
7.1	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
7.2	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
7.3	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002	0.0002
7.4	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
7.5	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
7.6	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
7.7	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
7.8	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0001
7.9	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
8.0	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
8.1	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
8.2	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
8.3	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
8.4	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
8.5	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
8.6	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
8.7	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
8.8	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
8.9	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0000
9.0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Bảng 2. Giá trị hàm  $\Phi_0(u) = \int_0^u \varphi(z) dz$ 

$$(\Phi_0(-u) = -\Phi_0(u))$$

u	0.00	.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706

1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.8	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.9	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
4.0	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000

Bảng 3. Giá trị hàm  $\Phi(u) = P(U < u)$ 

$$\Phi(u) = \int_{-\infty}^u \varphi(z) dz$$

u	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706

1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Bảng 4. Giá trị tới hạn phân phối student  $T_{\alpha}(n)$ 

$\alpha \backslash n$	0.2	0.15	0.1	0.05	0.025	0.01	0.005	0.0025	0.001
1	1.376	1.963	3.078	6.314	12.71	31.82	63.66	127.3	318.3
2	1.061	1.386	1.886	2.920	4.303	6.965	9.925	14.09	22.33
3	0.978	1.250	1.638	2.353	3.182	4.541	5.841	7.453	10.22
4	0.941	1.190	1.533	2.132	2.776	3.747	4.604	5.598	7.173
5	0.920	1.156	1.476	2.015	2.571	3.365	4.032	4.773	5.893
6	0.906	1.134	1.440	1.943	2.447	3.143	3.707	4.317	5.208
7	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.029	4.785
8	0.889	1.108	1.397	1.860	2.306	2.896	3.355	3.833	4.501
9	0.883	1.100	1.383	1.833	2.262	2.821	3.250	3.690	4.297
10	0.879	1.093	1.372	1.812	2.228	2.764	3.169	3.581	4.144
11	0.876	1.088	1.363	1.796	2.201	2.718	3.106	3.497	4.025
12	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.428	3.930
13	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.372	3.852
14	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.326	3.787
15	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.286	3.733
16	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.252	3.686
17	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.222	3.646
18	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.197	3.610
19	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.174	3.579
20	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.153	3.552
21	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.135	3.527
22	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.119	3.505

23	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.104	3.485
24	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.091	3.467
25	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.078	3.450
26	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.067	3.435
27	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.057	3.421
28	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.047	3.408
29	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.038	3.396
30	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.030	3.385
40	0.851	1.050	1.303	1.684	2.021	2.423	2.704	2.971	3.307
50	0.849	1.047	1.299	1.676	2.009	2.403	2.678	2.937	3.261
60	0.848	1.045	1.296	1.671	2.000	2.390	2.660	2.915	3.232
70	0.847	1.044	1.294	1.667	1.994	2.381	2.648	2.899	3.211
80	0.846	1.043	1.292	1.664	1.990	2.374	2.639	2.887	3.195
90	0.846	1.042	1.291	1.662	1.987	2.368	2.632	2.878	3.183
100	0.845	1.042	1.290	1.660	1.984	2.364	2.626	2.871	3.174
120	0.845	1.041	1.289	1.658	1.980	2.358	2.617	2.860	3.160
240	0.843	1.039	1.285	1.651	1.970	2.342	2.596	2.833	3.125
$\infty$	0.842	1.036	1.282	1.645	1.960	2.326	2.576	2.807	3.090

Bảng 5. Giá trị tới hạn phân phối khi-bình phương  $\chi^2_\alpha(n)$ 

$\alpha \backslash n$	0.995	0.99	0.975	0.95	0.9	0.1	0.05	0.025	0.01	0.005
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.60
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.34	12.84
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.14	13.28	14.86
5	0.412	0.554	0.831	1.145	1.610	9.236	11.07	12.83	15.09	16.75
6	0.676	0.872	1.237	1.635	2.204	10.64	12.59	14.45	16.81	18.55
7	0.989	1.239	1.690	2.167	2.833	12.02	14.07	16.01	18.48	20.28
8	1.344	1.646	2.180	2.733	3.490	13.36	15.51	17.53	20.09	21.95
9	1.735	2.088	2.700	3.325	4.168	14.68	16.92	19.02	21.67	23.59
10	2.156	2.558	3.247	3.940	4.865	15.99	18.31	20.48	23.21	25.19
11	2.603	3.053	3.816	4.575	5.578	17.28	19.68	21.92	24.72	26.76
12	3.074	3.571	4.404	5.226	6.304	18.55	21.03	23.34	26.22	28.30
13	3.565	4.107	5.009	5.892	7.042	19.81	22.36	24.74	27.69	29.82
14	4.075	4.660	5.629	6.571	7.790	21.06	23.68	26.12	29.14	31.32
15	4.601	5.229	6.262	7.261	8.547	22.31	25.00	27.49	30.58	32.80
16	5.142	5.812	6.908	7.962	9.312	23.54	26.30	28.85	32.00	34.27
17	5.697	6.408	7.564	8.672	10.09	24.77	27.59	30.19	33.41	35.72
18	6.265	7.015	8.231	9.390	10.86	25.99	28.87	31.53	34.81	37.16
19	6.844	7.633	8.907	10.12	11.65	27.20	30.14	32.85	36.19	38.58
20	7.434	8.260	9.591	10.85	12.44	28.41	31.41	34.17	37.57	40.00

21	8.034	8.897	10.28	11.59	13.24	29.62	32.67	35.48	38.93	41.40
22	8.643	9.542	10.98	12.34	14.04	30.81	33.92	36.78	40.29	42.80
23	9.260	10.20	11.69	13.09	14.85	32.01	35.17	38.08	41.64	44.18
24	9.886	10.86	12.40	13.85	15.66	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	34.38	37.65	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	35.56	38.89	41.92	45.64	48.29
27	11.81	12.88	14.57	16.15	18.11	36.74	40.11	43.19	46.96	49.64
28	12.46	13.56	15.31	16.93	18.94	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	51.81	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	37.69	63.17	67.50	71.42	76.15	79.49
60	35.53	37.48	40.48	43.19	46.46	74.40	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	55.33	85.53	90.53	95.02	100.4	104.2
80	51.17	53.54	57.15	60.39	64.28	96.58	101.9	106.6	112.3	116.3
90	59.20	61.75	65.65	69.13	73.29	107.6	113.1	118.1	124.1	128.3
100	67.33	70.06	74.22	77.93	82.36	118.5	124.3	129.6	135.8	140.2
120	83.85	86.92	91.57	95.70	100.6	140.2	146.6	152.2	159.0	163.6
150	109.1	112.7	118.0	122.7	128.3	172.6	179.6	185.8	193.2	198.4
200	152.2	156.4	162.7	168.3	174.8	226.0	234.0	241.1	249.4	255.3

Bảng 6. Giá trị tới hạn phân phối fisher  $F_{\alpha}(n_1; n_2)$

$n_2$	$\alpha$	$n_1$									
		1	2	3	4	5	6	7	8	9	10
5	0.1	4.060	3.780	3.619	3.520	3.453	3.405	3.368	3.339	3.316	3.297
	0.05	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735
	0.025	10.007	8.434	7.764	7.388	7.146	6.978	6.853	6.757	6.681	6.619
6	0.1	3.776	3.463	3.289	3.181	3.108	3.055	3.014	2.983	2.958	2.937
	0.05	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147	4.099	4.060
	0.025	8.813	7.260	6.599	6.227	5.988	5.820	5.695	5.600	5.523	5.461
7	0.1	3.589	3.257	3.074	2.961	2.883	2.827	2.785	2.752	2.725	2.703
	0.05	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.677	3.637
	0.025	8.073	6.542	5.890	5.523	5.285	5.119	4.995	4.899	4.823	4.761
8	0.1	3.458	3.113	2.924	2.806	2.726	2.668	2.624	2.589	2.561	2.538
	0.05	5.318	4.459	4.066	3.838	3.687	3.581	3.500	3.438	3.388	3.347
	0.025	7.571	6.059	5.416	5.053	4.817	4.652	4.529	4.433	4.357	4.295
9	0.1	3.360	3.006	2.813	2.693	2.611	2.551	2.505	2.469	2.440	2.416
	0.05	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230	3.179	3.137
	0.025	7.209	5.715	5.078	4.718	4.484	4.320	4.197	4.102	4.026	3.964
10	0.1	3.285	2.924	2.728	2.605	2.522	2.461	2.414	2.377	2.347	2.323
	0.05	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020	2.978
	0.025	6.937	5.456	4.826	4.468	4.236	4.072	3.950	3.855	3.779	3.717
11	0.1	3.225	2.860	2.660	2.536	2.451	2.389	2.342	2.304	2.274	2.248
	0.05	4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948	2.896	2.854

	0.025	6.724	5.256	4.630	4.275	4.044	3.881	3.759	3.664	3.588	3.526
12	0.1	3.177	2.807	2.606	2.480	2.394	2.331	2.283	2.245	2.214	2.188
	0.05	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849	2.796	2.753
	0.025	6.554	5.096	4.474	4.121	3.891	3.728	3.607	3.512	3.436	3.374
13	0.1	3.136	2.763	2.560	2.434	2.347	2.283	2.234	2.195	2.164	2.138
	0.05	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767	2.714	2.671
	0.025	6.414	4.965	4.347	3.996	3.767	3.604	3.483	3.388	3.312	3.250
14	0.1	3.102	2.726	2.522	2.395	2.307	2.243	2.193	2.154	2.122	2.095
	0.05	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699	2.646	2.602
	0.025	6.298	4.857	4.242	3.892	3.663	3.501	3.380	3.285	3.209	3.147
15	0.1	3.073	2.695	2.490	2.361	2.273	2.208	2.158	2.119	2.086	2.059
	0.05	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588	2.544
	0.025	6.200	4.765	4.153	3.804	3.576	3.415	3.293	3.199	3.123	3.060
16	0.1	3.048	2.668	2.462	2.333	2.244	2.178	2.128	2.088	2.055	2.028
	0.05	4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591	2.538	2.494
	0.025	6.115	4.687	4.077	3.729	3.502	3.341	3.219	3.125	3.049	2.986
17	0.1	3.026	2.645	2.437	2.308	2.218	2.152	2.102	2.061	2.028	2.001
	0.05	4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548	2.494	2.450
	0.025	6.042	4.619	4.011	3.665	3.438	3.277	3.156	3.061	2.985	2.922
18	0.1	3.007	2.624	2.416	2.286	2.196	2.130	2.079	2.038	2.005	1.977
	0.05	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510	2.456	2.412
	0.025	5.978	4.560	3.954	3.608	3.382	3.221	3.100	3.005	2.929	2.866
19	0.1	2.990	2.606	2.397	2.266	2.176	2.109	2.058	2.017	1.984	1.956
	0.05	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477	2.423	2.378
	0.025	5.922	4.508	3.903	3.559	3.333	3.172	3.051	2.956	2.880	2.817

Bảng 6. Giá trị tới hạn phân phối fisher  $F_{\alpha}(n_1; n_2)$   
(tiếp theo)

n <sub>2</sub>	α	n <sub>1</sub>									
		1	2	3	4	5	6	7	8	9	10
20	0.1	2.975	2.589	2.380	2.249	2.158	2.091	2.040	1.999	1.965	1.937
	0.05	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447	2.393	2.348
	0.025	5.871	4.461	3.859	3.515	3.289	3.128	3.007	2.913	2.837	2.774
21	0.1	2.961	2.575	2.365	2.233	2.142	2.075	2.023	1.982	1.948	1.920
	0.05	4.325	3.467	3.072	2.840	2.685	2.573	2.488	2.420	2.366	2.321
	0.025	5.827	4.420	3.819	3.475	3.250	3.090	2.969	2.874	2.798	2.735
22	0.1	2.949	2.561	2.351	2.219	2.128	2.060	2.008	1.967	1.933	1.904
	0.05	4.301	3.443	3.049	2.817	2.661	2.549	2.464	2.397	2.342	2.297
	0.025	5.786	4.383	3.783	3.440	3.215	3.055	2.934	2.839	2.763	2.700
25	0.1	2.918	2.528	2.317	2.184	2.092	2.024	1.971	1.929	1.895	1.866
	0.05	4.242	3.385	2.991	2.759	2.603	2.490	2.405	2.337	2.282	2.236
	0.025	5.686	4.291	3.694	3.353	3.129	2.969	2.848	2.753	2.677	2.613
30	0.1	2.881	2.489	2.276	2.142	2.049	1.980	1.927	1.884	1.849	1.819
	0.05	4.171	3.316	2.922	2.690	2.534	2.421	2.334	2.266	2.211	2.165
	0.025	5.568	4.182	3.589	3.250	3.026	2.867	2.746	2.651	2.575	2.511
40	0.1	2.835	2.440	2.226	2.091	1.997	1.927	1.873	1.829	1.793	1.763
	0.05	4.085	3.232	2.839	2.606	2.449	2.336	2.249	2.180	2.124	2.077
	0.025	5.424	4.051	3.463	3.126	2.904	2.744	2.624	2.529	2.452	2.388
50	0.1	2.809	2.412	2.197	2.061	1.966	1.895	1.840	1.796	1.760	1.729
	0.05	4.034	3.183	2.790	2.557	2.400	2.286	2.199	2.130	2.073	2.026
	0.025	5.340	3.975	3.390	3.054	2.833	2.674	2.553	2.458	2.381	2.317

**Bảng 7: Giá trị tới hạn durbin-watson mức 5%**

n	k'	1		2		3		4		5	
		dL	dU	dL	dU	dL	dU	dL	dU	dL	dU
6		0.610	1.40	0.467	1.896	0.368	2.287	0.296	2.588		
7		0.700	1.356	0.559	1.777	0.445	2.128				
8		0.763	1.332	0.629	1.699						
9		0.924	1.320								
10		0.879	1.320	0.697	1.641	0.525	2.016	0.376	2.414	0.243	2.882
11		0.927	1.324	0.658	1.604	0.595	1.928	0.444	2.283	0.316	2.645
12		0.971	1.331	0.812	1.579	0.658	1.834	0.512	2.177	0.379	2.506
13		1.010	1.340	0.861	1.562	0.715	1.816	0.574	2.094	0.445	2.390
14		1.045	1.350	0.605	1.551	0.767	1.779	0.632	2.030	0.505	2.296
15		1.077	1.361	0.946	1.543	0.814	1.750	0.685	1.977	0.562	2.220
16		1.106	1.371	0.982	1.539	0.857	1.728	0.734	1.935	0.615	2.157
17		1.133	1.381	1.015	1.536	0.897	1.710	0.779	1.900	0.664	2.104
18		1.158	1.391	1.046	1.535	0.933	1.696	0.820	1.872	0.710	2.060
19		1.180	1.401	1.074	1.536	0.967	1.685	0.859	1.848	0.752	2.020
20		1.201	1.411	1.100	1.537	0.998	1.676	0.894	1.828	0.792	1.991
21		1.221	1.420	1.125	1.538	1.026	1.699	0.927	1.812	0.829	1.964
22		1.239	1.429	1.147	1.541	1.053	1.664	0.958	1.797	0.863	1.940
23		1.257	1.437	1.168	1.543	1.078	1.660	0.986	1.785	0.859	1.920
24		1.273	1.446	1.188	1.546	1.101	1.656	1.103	1.775	0.925	1.902
25		1.288	1.454	1.206	1.550	1.123	1.654	1.038	1.767	0.953	1.886
26		1.302	1.461	1.224	1.553	1.143	1.652	1.062	1.759	0.979	1.873
27		1.316	1.496	1.240	1.556	1.162	1.651	1.084	1.853	1.004	1.861
28		1.328	1.476	1.255	1.560	1.181	1.650	1.104	1.747	1.028	1.850
29		1.341	1.483	1.270	1.563	1.198	1.650	1.124	1.734	1.050	1.841
30		1.352	1.489	1.281	1.567	1.214	1.650	1.143	1.739	1.071	1.833
35		1.402	1.519	1.343	1.584	1.283	1.653	1.222	1.762	1.160	1.803
40		1.442	1.544	1.391	1.600	1.338	1.659	1.285	1.821	1.230	1.786
45		1.475	1.566	1.430	1.615	1.383	1.666	1.336	1.720	1.287	1.776
50		1.503	1.585	1.462	1.628	1.421	1.674	1.378	1.721	1.335	1.771
55		1.528	1.601	1.490	1.641	1.452	1.681	1.414	1.724	1.374	1.768
60		1.549	1.616	1.514	1.652	1.480	1.689	1.444	1.727	1.108	1.767
65		1.567	1.629	1.536	1.662	1.503	1.696	1.471	1.731	1.438	1.737
70		1.583	1.641	1.554	1.672	1.525	1.703	1.494	1.735	1.461	1.768
75		1.598	1.652	1.571	1.680	1.543	1.709	1.515	1.739	1.487	1.770
80		1.611	1.662	1.586	1.688	1.560	1.715	1.534	1.743	1.507	1.772
85		1.624	1.671	1.600	1.696	1.575	1.721	1.550	1.747	1.525	1.774
90		1.635	1.679	1.612	1.703	1.589	1.726	1.566	1.751	1.542	1.776
100		1.654	1.694	1.634	1.715	1.613	1.736	1.592	1.788	1.571	1.780
150		1.720	1.746	1.706	1.760	1.693	1.774	1.679	1.788	1.665	1.802
200		1.758	1.778	1.748	1.789	1.738	1.799	1.728	1.810	1.718	1.820

60	0.1	2.791	2.393	2.177	2.041	1.946	1.875	1.819	1.775	1.738	1.707
	0.05	4.001	3.150	2.758	2.525	2.368	2.254	2.167	2.097	2.040	1.993
	0.025	5.286	3.925	3.343	3.008	2.786	2.627	2.507	2.412	2.334	2.270
70	0.1	2.779	2.380	2.164	2.027	1.931	1.860	1.804	1.760	1.723	1.691
	0.05	3.978	3.128	2.736	2.503	2.346	2.231	2.143	2.074	2.017	1.969
	0.025	5.247	3.890	3.309	2.975	2.754	2.595	2.474	2.379	2.302	2.237
80	0.1	2.769	2.370	2.154	2.016	1.921	1.849	1.793	1.748	1.711	1.680
	0.05	3.960	3.111	2.719	2.486	2.329	2.214	2.126	2.056	1.999	1.951
	0.025	5.218	3.864	3.284	2.950	2.730	2.571	2.450	2.355	2.277	2.213
90	0.1	2.762	2.363	2.146	2.008	1.912	1.841	1.785	1.739	1.702	1.670
	0.05	3.947	3.098	2.706	2.473	2.316	2.201	2.113	2.043	1.986	1.938
	0.025	5.196	3.844	3.265	2.932	2.711	2.552	2.432	2.336	2.259	2.194
100	0.1	2.756	2.356	2.139	2.002	1.906	1.834	1.778	1.732	1.695	1.663
	0.05	3.936	3.087	2.696	2.463	2.305	2.191	2.103	2.032	1.975	1.927
	0.025	5.179	3.828	3.250	2.917	2.696	2.537	2.417	2.321	2.244	2.179
120	0.1	2.748	2.347	2.130	1.992	1.896	1.824	1.767	1.722	1.684	1.652
	0.05	3.920	3.072	2.680	2.447	2.290	2.175	2.087	2.016	1.959	1.910
	0.025	5.152	3.805	3.227	2.894	2.674	2.515	2.395	2.299	2.222	2.157
150	0.1	2.739	2.338	2.121	1.983	1.886	1.814	1.757	1.712	1.674	1.642
	0.05	3.904	3.056	2.665	2.432	2.274	2.160	2.071	2.001	1.943	1.894
	0.025	5.126	3.781	3.204	2.872	2.652	2.494	2.373	2.278	2.200	2.135
180	0.1	5.109	3.766	3.189	2.858	2.638	2.479	2.359	2.263	2.185	2.120
	0.05	6.778	4.725	3.892	3.425	3.120	2.904	2.740	2.611	2.507	2.421
	0.025	8.077	5.457	4.423	3.851	3.481	3.219	3.022	2.869	2.744	2.642

**Bảng 7: Giá trị tới hạn durbin-watson mức 5%**

k' n	6		7		8		9		10	
	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU
6										
7										
8										
9										
10	0.203	3.005	0.171	3.149	0.147	3.266	0.127	3.360		
11	0.268	2.832	0.230	2.985	0.200	3.111				
12	0.328	2.692	0.286	2.848						
13	0.389	2.572								
14										
15	0.447	2.472	0.343	2.727	0.251	2.979	0.175	3.216	0.111	3.438
16	0.502	2.388	0.398	2.624	0.304	2.860	0.222	3.090	0.155	3.304
17	0.554	2.318	0.451	2.537	0.356	2.757	0.242	2.975	0.198	3.184
18	0.603	2.257	0.502	2.461	0.407	2.667	0.321	2.873	0.244	3.073
19	0.649	2.206	0.549	2.396	0.456	2.589	0.369	2.783	0.290	2.974
20	0.692	2.162	0.595	2.339	0.502	2.521	0.416	2.704	0.336	2.885
21	0.732	2.124	0.637	2.290	0.547	2.460	0.461	2.633	0.380	2.806
22	0.769	2.090	0.677	2.246	0.588	2.407	0.504	2.571	0.424	2.734
23	0.804	2.061	0.715	2.208	0.628	2.360	0.545	2.514	0.465	2.670
24	0.837	2.035	0.751	2.174	0.666	2.318	0.581	2.464	0.506	2.613
25	0.868	2.012	0.784	2.144	0.702	2.280	0.621	2.419	0.544	2.560
26	0.897	1.992	0.816	2.117	0.735	2.246	0.657	2.379	0.581	2.513
27	0.925	1.974	0.845	2.093	0.767	2.216	0.691	2.342	0.616	2.470
28	0.951	1.958	0.874	2.071	0.798	2.188	0.823	2.309	0.650	2.431
29	0.975	1.944	0.900	2.052	0.826	2.164	0.853	2.278	0.682	2.396
30	0.998	1.931	0.926	2.034	0.854	2.141	0.872	2.251	0.712	2.363
35	1.097	1.884	1.034	1.967	0.971	2.054	0.908	2.144	0.845	2.236
40	1.175	1.854	1.120	1.924	1.064	1.997	1.008	2.072	0.952	2.149
45	1.238	1.835	1.189	1.895	1.139	1.958	1.089	2.022	1.038	2.088
50	1.291	1.822	1.246	1.875	1.201	1.930	1.156	1.983	1.110	2.044
55	1.334	1.814	1.294	1.861	1.253	1.909	1.212	1.959	1.170	2.010
60	1.372	1.808	1.335	1.850	1.298	1.260	1.260	1.939	1.222	1.984
65	1.404	1.805	1.370	1.843	1.336	1.882	1.301	1.923	1.266	1.964
70	1.433	1.802	1.401	1.837	1.369	1.873	1.337	1.910	1.305	1.948
75	1.458	1.801	1.428	1.834	1.399	1.867	1.369	1.901	1.339	1.935
80	1.480	1.801	1.453	1.831	1.425	1.861	1.397	1.893	1.369	1.925
85	1.500	1.801	1.474	1.829	1.448	1.857	1.422	1.886	1.396	1.916
90	1.518	1.801	1.494	1.827	1.469	1.854	1.445	1.881	1.420	1.909
100	1.550	1.803	1.528	1.826	1.506	1.884	1.484	1.874	1.462	1.898
150	1.651	1.817	1.637	1.832	1.622	1.847	1.608	1.862	1.594	1.877
200	1.707	1.831	1.697	1.841	1.686	1.852	1.675	1.863	1.665	1.874

## TÀI LIỆU THAM KHẢO

1. Phạm Trí Cao – Vũ Minh Châu (2014), *Kinh tế lượng ứng dụng*, Nhà xuất bản Lao động xã hội.
2. Nguyễn Quang Đông - Nguyễn Thị Minh (2012), *Giáo trình Kinh tế lượng*, Nhà xuất bản Đại học kinh tế quốc dân.
3. Hoàng Ngọc Nhậm (chủ biên) (2009), *Giáo trình kinh tế lượng*, Nhà xuất bản Lao động xã hội.
4. Hoàng Ngọc Nhậm (chủ biên) (2009), *Bài tập kinh tế lượng với sự trợ giúp của Eviews*, Nhà xuất bản Lao động xã hội.
5. Damodar N Gujarati (2004), *Basic Econometrics*, Fourth Edition, McGraw Hill.