

TRƯỜNG ĐẠI HỌC TÀI CHÍNH - MARKETING
BỘ MÔN TOÁN – THỐNG KÊ

Slide bài giảng và bài tập

MÔN KINH TẾ LƯỢNG
(Econometric)

Giảng viên : ThS. Nguyễn Trung Đông

Tp. Hồ Chí Minh, 02 - 02 - 2014

TRƯỜNG ĐẠI HỌC
TÀI CHÍNH - MARKETING
KHOA CƠ BẢN

Môn : **KINH TẾ LƯỢNG**
(Econometric)

Số tín chỉ : 3
Số tiết : 30 LT + 30 TH

Giảng viên : ThS. Nguyễn Trung Đông

1

TRƯỜNG ĐẠI HỌC
TÀI CHÍNH - MARKETING
KHOA CƠ BẢN

Môn : **KINH TẾ LƯỢNG**
(Econometric)

Hình thức đánh giá môn học

Điểm quá trình (30%)

Điểm kết thúc học (70%)

Điểm học phần = (Điểm quá trình + Điểm kết thúc học)

Giảng viên : ThS. Nguyễn Trung Đông

TỔNG QUAN

Mục tiêu môn học:

- Cung cấp phương pháp phân tích định lượng.
- Ứng dụng: Phương pháp định lượng
- ❖ Làm thực tập tốt nghiệp, luận văn tốt nghiệp.
- ❖ Phân tích, kiểm định và dự báo kinh tế.

3

NỘI DUNG MÔN HỌC

- ❖ Ôn tập
- ❖ Chương 0. Mở đầu
- ❖ Chương 1. Hồi quy hai biến
- ❖ Chương 2. Hồi quy bội
- ❖ Chương 3. Kiểm định giả thuyết mô hình

4

TÀI LIỆU THAM KHẢO

Tiếng Việt

- 1) Đinh Ngọc Thanh, Nguyễn Văn Phong, Nguyễn Trung Đông, Nguyễn Thị Hải Ninh: *Giáo trình kinh tế lượng*, lưu hành nội bộ, Đại học tài chính – Marketing.
- 2) Phạm Chí Cao – Vũ Minh Châu: *Kinh tế lượng ứng dụng*, nhà xuất bản Thống kê, 2010.
- 3) Nguyễn Quang Đông: *Bài giảng Kinh tế lượng*, nhà xuất bản thống kê, 2006.
- 4) Chương trình giảng dạy Kinh tế Fullbright: *Bài giảng Kinh tế lượng*, 2004.
- 5) Huỳnh Đạt Hùng, Nguyễn Khánh Bình, Phạm Xuân Giang: *Kinh tế lượng*, nhà xuất bản Phương Đông, 2012.

6) Nguyễn Cao Văn – Bùi Dương Hải, *Kinh tế lượng* (hướng dẫn và trả lời lý thuyết và bài tập, nhà xuất bản Tài Chính.

7) Bùi Minh trí: *Kinh tế lượng*, nhà xuất bản khoa học và kỹ thuật, 2006.

Tiếng Anh

- 1) Dimitrios Asteriou and Stephen G. Hall: *Applied Econometrics*, Published by Palgrave Macmillan, 2007.
- 2) Christopher Dougherty: *Introduction to Econometrics*, Published Oxford.
- 3) Jeffrey M. Wooldridge: *Introduction to Econometrics,...*
- 4) Damodar N Gujarati, *Basic Econometrics*, Mc Graw – Hill Inc, third edition, 1995.

6

Chương 0.**Ôn Tập**

Kinh tế lượng (Econometric): Lượng hóa các vấn đề về kinh tế.

1. Đạo hàm (tỷ lệ sự thay đổi)

Xét hàm số $Y=f(X)$. Trong đó

Y : Biến phụ thuộc, biến được giải thích, biến nội sinh, biến hồi quy.

X : Biến độc lập, biến giải thích, biến ngoại sinh.

Ví dụ 1: Thu nhập (X) – Chi tiêu (Y).

Lạm phát (X) – Lãi suất (Y).

2. Đạo hàm tại điểm. Xét hàm số: $y = f(x)$

$$\frac{f(x) - f(a)}{x - a} = \frac{\Delta y}{\Delta x} \quad \begin{array}{l} \Delta y : \text{sự thay đổi của } y \\ \Delta x : \text{sự thay đổi của } x \end{array}$$

Sự thay đổi của y theo x : $y' = f'(a) \approx \Delta y / \Delta x$

Tỷ lệ sự thay đổi của y theo x xung quanh điểm a .

Ví dụ 2: Xét mối quan hệ: $y = f(x)$

Giả sử: x : lạm phát, y : lãi suất và $f'(5) = 1.25$

Ý nghĩa: Nếu LP tăng 1% thì LS tăng 1.25%.

3. Đạo hàm riêng. Xét hàm số: $z = f(x, y)$

z : là biến phụ thuộc (biến được giải thích)

x, y : là biến độc lập (biến giải thích)

3.1. Đạo hàm riêng của z theo x

$$\frac{\partial z}{\partial x} = \lim_{\Delta x \rightarrow 0} \left(\frac{f(x + \Delta x, y) - f(x, y)}{\Delta x} \right); (\Delta x, \Delta y = 0)$$

3.2. Đạo hàm riêng của z theo y

$$\frac{\partial z}{\partial y} = \lim_{\Delta y \rightarrow 0} \left(\frac{f(x, y + \Delta y) - f(x, y)}{\Delta y} \right); (\Delta x = 0, \Delta y)$$

Ví dụ 3: $\frac{\partial z}{\partial x}(3, 2) = 0.4$; $\frac{\partial z}{\partial y}(3, 2) = 0.1$

Ví dụ 4: Tính đạo hàm riêng của hàm số sau

$$1. f(x, y) = x^3 + y^3 - 6xy + 2x + 3y + 1$$

$$2. f(x, y) = \ln(x^2 + y^2)$$

Giải

$$1. \frac{\partial f}{\partial x}(x, y) = 3x^2 - 6y + 2; \frac{\partial f}{\partial y}(x, y) = 3y^2 - 6x + 3$$

$$2. \frac{\partial f}{\partial x}(x, y) = \frac{2x}{x^2 + y^2}; \frac{\partial f}{\partial y}(x, y) = \frac{2y}{x^2 + y^2}$$

4. Điều kiện cần của cực trị.

Xét hàm số: $z = f(x, y)$

Hàm số đạt cực trị tại (x_0, y_0)

$$\Rightarrow \begin{cases} \frac{\partial f}{\partial x}(x_0, y_0) = 0 \\ \frac{\partial f}{\partial y}(x_0, y_0) = 0 \end{cases} \quad (*)$$

Nếu (x_0, y_0) thỏa (*) thì (x_0, y_0) được gọi là điểm dừng.

5. Điều kiện đủ của cực trị.

Xét điểm dừng: (x_0, y_0) . Đặt $A = \frac{\partial^2 f}{\partial x^2}(x_0, y_0)$,

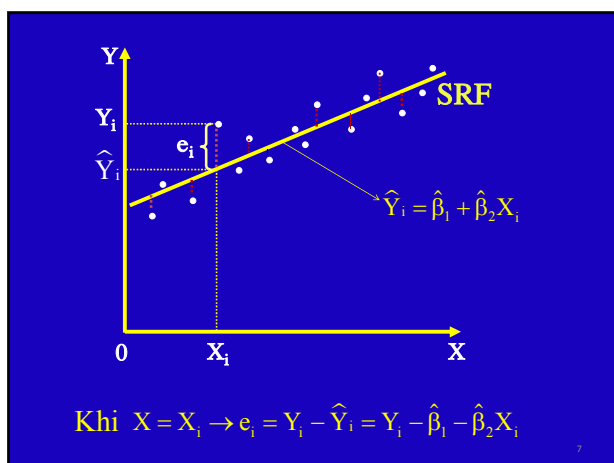
$$C = \frac{\partial^2 f}{\partial^2 y}(x_0, y_0), B = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0), \Delta = AC - B^2$$

Th1: Nếu $\Delta > 0$ và $A > 0$ thì (x_0, y_0) là cực tiểu.

Th2: Nếu $\Delta > 0$ và $A < 0$ thì (x_0, y_0) là cực đại.

Th3: Nếu $\Delta < 0$ thì (x_0, y_0) không là cực trị.

Th4: Nếu $\Delta = 0$ chưa đủ cơ sở kết luận.

**Phương pháp bình phương cực tiểu (OLS : Ordinary least squares)**

Tổng bình phương các sai lệch

(RSS : Residual sum of squares)

$$RSS = e_1^2 + e_2^2 + \dots + e_n^2 = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2$$

Bài toán. Tìm $(\hat{\beta}_1, \hat{\beta}_2)$ sao cho $RSS \rightarrow \min$

$$\frac{\partial RSS}{\partial \hat{\beta}_1}(\hat{\beta}_1, \hat{\beta}_2) = \sum_{i=1}^n 2(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)(-1) = 0$$

$$\frac{\partial RSS}{\partial \hat{\beta}_2}(\hat{\beta}_1, \hat{\beta}_2) = \sum_{i=1}^n 2(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)(-X_i) = 0$$

Suy ra

$$\begin{cases} n\hat{\beta}_1 + \hat{\beta}_2 \sum_{i=1}^n X_i = \sum_{i=1}^n Y_i & \boxed{1} \\ \hat{\beta}_1 \sum_{i=1}^n X_i + \hat{\beta}_2 \sum_{i=1}^n X_i^2 = \sum_{i=1}^n X_i Y_i & \boxed{2} \end{cases}$$

Hệ Cramer

$$\begin{vmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{vmatrix} = n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2 > 0$$

$$(1) \Leftrightarrow \hat{\beta}_1 + \hat{\beta}_2 \frac{1}{n} \sum_{i=1}^n X_i = \frac{1}{n} \sum_{i=1}^n Y_i$$

$$\Leftrightarrow \hat{\beta}_1 + \hat{\beta}_2 \bar{X} = \bar{Y} \Leftrightarrow \hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}$$

$$\hat{\beta}_2 = \frac{\begin{vmatrix} n & \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i Y_i \end{vmatrix}}{\begin{vmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{vmatrix}} = \frac{n \sum_{i=1}^n X_i Y_i - \left(\sum_{i=1}^n X_i \right) \left(\sum_{i=1}^n Y_i \right)}{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2}$$

Ví dụ:

X	1	2	3	4	5
Y	2	5	7	8	9

6. Phân phối xác suất

6.1. Phân phối chuẩn

$X \sim N(\mu, \sigma^2)$

$$P(a \leq X \leq b) = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{t^2}{2}} dt$$

Đặt $t = \frac{x-\mu}{\sigma}$

Nếu $X \sim N(\mu, \sigma^2)$, đặt $Y = \frac{X-\mu}{\sigma}$ thì $Y \sim N(0,1)$

$$P(a \leq X \leq b) = P\left(\frac{a-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right) = P\left(\frac{a-\mu}{\sigma} \leq Y \leq \frac{b-\mu}{\sigma}\right)$$

Bài toán cho $Y \sim N(0,1), \alpha \leq \beta$

Tìm $P(\alpha \leq Y \leq \beta) = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{t^2}{2}} dt$

$$= \frac{1}{\sqrt{2\pi}} \int_0^{\beta} e^{-\frac{t^2}{2}} dt - \frac{1}{\sqrt{2\pi}} \int_0^{\alpha} e^{-\frac{t^2}{2}} dt = \Phi(\beta) - \Phi(\alpha)$$

Trong đó

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt \quad \text{Laplace}$$

Lấy $x = 0.00, 0.01, \dots, 3.99$ suy ra bảng phân phối Gauss

Ví dụ : $\Phi(1.26) = 0.3962$

Với $x \geq 4, \Phi(x) = 0.5$

Nếu $x < 0, \Phi(-x) = -\Phi(x)$

6.2. Phân phối Student St(n)**a) Một số kết quả**

- i) Nếu $X \sim N(0,1)$ thì $X^2 \sim \chi^2(1)$
- ii) Nếu X, Y độc lập, $X \sim \chi^2(n)$; $Y \sim \chi^2(m)$ thì
 $X + Y \sim \chi^2(n+m)$
- iii) Cho $X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$ và độc lập

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

$$S_X^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \mu_X)^2 \quad (\text{Phương sai có hiệu chỉnh})$$

$$S_X^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu_X)^2 \quad (\text{Phương sai không hiệu chỉnh})$$

13

b) Định nghĩa phân phối Student

Nếu $X \sim N(0,1)$; $Y \sim \chi^2(n)$ và X, Y độc lập thì

$$T = \frac{X}{\sqrt{Y/n}} \sim \text{St}(n)$$

c) Định lý Lindeberg – Levy

Cho $X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$

i) $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

ii) $\frac{(n-1)S_X^2}{\sigma^2} \sim \chi^2(n-1)$

14

Trong đó \bar{X}, S_X^2 lần lượt trung bình và phương sai mẫu có hiệu chỉnh

Chú ý :

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow Y = \frac{(\bar{X} - \mu)\sqrt{n}}{\sigma} \sim N(0,1)$$

$$Z = \frac{(n-1)S_X^2}{\sigma^2} \sim \chi^2(n-1)$$

$$T = \frac{Y}{\sqrt{Z/n-1}} = \frac{(\bar{X} - \mu)\sqrt{n}}{S_X} \sim \text{St}(n-1)$$

15

6.3. Phân phối Fisher

Nếu

$$X \sim \chi^2(n), Y \sim \chi^2(m)$$

và X, Y độc lập thì

$$F = \frac{X/n}{Y/m} \sim F(n, m)$$

16

7. Tìm khoảng tin cậy

Gọi $[a, b]$ là khoảng tin cậy (KTC) với độ tin cậy γ

Định nghĩa: $P(a \leq X \leq b) = \gamma \rightarrow 0.9, 0.95, 0.99$

Nguy cơ sai lầm $\alpha = 1 - \gamma$

7.1. $X \sim N(0, 1)$

Chọn KTC cho X là $[-C, C]$ sao cho

$$P(-C \leq X \leq C) = \gamma \Leftrightarrow \Phi(C) = \frac{\gamma}{2}$$

Ký hiệu:

$$C = Z_{\gamma/2}$$

17

7.2. $X \sim N(\mu, \sigma^2)$. Đặt $Y = \frac{X - \mu}{\sigma}$ thì $Y \sim N(0, 1)$

Chọn KTC cho Y là $[-C, C]$ sao cho

$$P(-C \leq Y \leq C) = \gamma$$

Khoảng tin cậy cho X : $X \in [\mu - C\sigma; \mu + C\sigma]$

7.3. $T \sim St(n)$

Chọn KTC là cho T là $[-C, C]$ sao cho

$$P(-C \leq T \leq C) = \gamma$$

Với $C = t_{\alpha}^n$

Chú ý: khi $n \geq 30$ thì $St(n) \equiv N(0, 1)$

18

7.4. $F \sim F(n, m)$

Chọn KTC cho F là $[0, C]$ sao cho $P(0 \leq F \leq C) = \gamma$

Với $C = f_{\alpha}(n, m)$

7.5. $X \sim \chi^2(n)$

Chọn KTC cho X

- Dạng $[a, b]$ sao cho $P(a \leq X \leq b) = \gamma$

Với $a = \chi_{1-\frac{\alpha}{2}}^2(n)$; $b = \chi_{\frac{\alpha}{2}}^2(n)$

- Dạng $[0, C]$ sao cho $P(0 \leq X \leq C) = \gamma$

Với $C = \chi_{\alpha}^2(n)$

..

Bài Giảng
KINH TẾ LƯỢNG
 (Econometric)

Chương Mở Đầu

GV: ThS. Nguyễn Trung Đông
 nguyentrungdong144@yahoo.com

Chương Mở Đầu

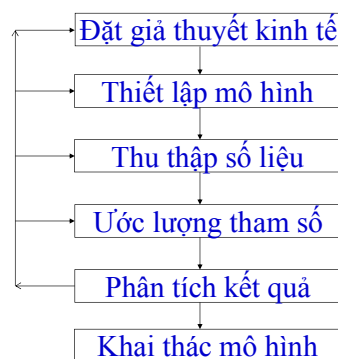
- ❖ Khái niệm về kinh tế lượng.
- ❖ Phương pháp luận kinh tế lượng.
- ❖ Đánh giá sơ bộ số liệu thống kê.

2

1. Khái niệm về kinh tế lượng

- ❖ Econometric= Econo + Metric
- ❖ Khái niệm: Kinh tế lượng là nghiên cứu những vấn đề thực nghiệm của các quy luật kinh tế; thông qua việc xây dựng, phân tích, đánh giá các mô hình cho ra lời giả bằng số, hỗ trợ việc ra quyết định.
- ❖ Kinh tế lượng sử dụng các kết quả của:
 - + Mô hình toán kinh tế; Xác suất và thống kê toán; Toán cao cấp, Lý thuyết kinh tế.
 - + Một phần mềm thông dụng: Eview, SPSS,...

2. Phương pháp luận của kinh tế lượng



4

2. Phương pháp luận của kinh tế lượng

Khi mô hình nhận được phù hợp với giả thuyết kinh tế, ta có thể dùng mô hình này để :

- ❖ Dự báo kết quả
- ❖ Kiểm tra hay đề ra chính sách

5

Phân tích tác động của thu nhập lên tiêu dùng tại các quốc gia vùng Đông Á – Thái Bình Dương năm 1998

Bước 1. Đặt giả thuyết kinh tế

Theo Keynes: Con người thường tăng tiêu dùng khi thu nhập của họ tăng lên, nhưng không nhiều như gia tăng thu nhập của họ.

6

Phân tích tác động của thu nhập lên tiêu dùng tại các quốc gia vùng Đông Á – Thái Bình Dương năm 1998

Bước 2. Thiết lập mô hình toán

Dạng đơn giản thể hiện mối quan hệ giữa thu nhập (TN) và tiêu dùng (TD) là dạng tuyến tính

$$TD = \beta_1 + \beta_2 TN$$

Trong đó β_1, β_2 là hai tham số và $0 < \beta_2 < 1$

7

Phân tích tác động của thu nhập lên tiêu dùng tại các quốc gia vùng Đông Á – Thái Bình Dương năm 1998

Bước 3. Thu thập số liệu

DVT: tỷ USD

Quốc gia	Tiêu dùng	Thu nhập	Quốc gia	Tiêu dùng	Thu nhập
Australia	289.35	372.72	Macao	3.3185	6.4474
Cambodia	2.7132	2.8709	Malaysia	37.344	72.488
China	560.53	946.31	Mongolia	0.76041	1.0417
Fiji	1.3677	1.5774	New Zealand	42.507	52.944
Hong Kong	113.88	162.94	Papua New Guinea	2.9644	3.8208
Indonesia	62.779	98.827	Philippines	57.088	65.535
Japan	2715.3	3808.1	Singapore	40.911	82.773
Korea, Rep.	208.48	317.08	Thailand	73.261	112.09
Lao PDR	0.94699	1.2609	Vietnam	21.443	27.184

Nguồn: World Development Indicators 2001, WB.

Phân tích tác động của thu nhập lên tiêu dùng tại các quốc gia vùng Đông Á – Thái Bình Dương năm 1998

Bước 4. Ước lượng tham số

Để ước lượng các tham số, ta sử dụng phương pháp bình phương cực tiểu (OLS)

$$TD = -6,28 + 0,709TN$$

9

Phân tích tác động của thu nhập lên tiêu dùng tại các quốc gia vùng Đông Á – Thái Bình Dương năm 1998

Bước 5. Phân tích kết quả

Do hệ số $\beta_2 = 0,709$ thỏa điều kiện

$0 < \beta_2 < 1$ nên kết quả này phù hợp với giả thuyết kinh tế trong bước 1.

10

Phân tích tác động của thu nhập lên tiêu dùng tại các quốc gia vùng Đông Á – Thái Bình Dương năm 1998

Bước 6. Khai thác kết quả

Ta có thể dùng mô hình ở bước 4 để dự báo. Chẳng hạn, nếu biết thu nhập là 300 tỷ USD thì dự đoán tiêu dùng trung bình là $TD = -6,28 + 0,709 \cdot 300 = 206,43$ tỷ USD.

11

3. Đánh giá về số liệu thống kê

Khi đã có số liệu thống kê

- ❖ Ta có thể biểu diễn chúng bằng biểu đồ rời rạc (phân tán).
- ❖ Ước lượng mối quan hệ giữa chúng bằng một số các tham số thống kê như: hiệp phương sai, hệ số tương quan.

12

3. Đánh giá về số liệu thống kê

❖ Hiệp phương sai tổng thể

$$\begin{aligned}\sigma_{X,Y} &= E[(X - \mu_X)(Y - \mu_Y)] \\ &= \frac{1}{N} \sum_{i=1}^N (X_i - \mu_X)(Y_i - \mu_Y)\end{aligned}$$

❖ Hệ số tương quan tổng thể

$$\rho_{X,Y} = \frac{\sigma_{X,Y}}{\sigma_X \sigma_Y}$$

13

3. Đánh giá về số liệu thống kê

❖ Hiệp phương sai mẫu

$$\begin{aligned}S_{X,Y} &= E[(X - \bar{X})(Y - \bar{Y})] \\ &= \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})\end{aligned}$$

❖ Hệ số tương quan mẫu

$$r_{X,Y} = \frac{S_{X,Y}}{S_X S_Y}$$

14

Từ đó, ta có đánh giá sơ bộ số liệu sau

(i) Khi $r_{X,Y} = 0$, ta nói X và Y *độc lập* (hay có *quan hệ phi tuyến* với nhau).

(ii) Khi $|r_{X,Y}| \approx 1$, ta nói X và Y có *quan hệ chặt chẽ* với nhau.

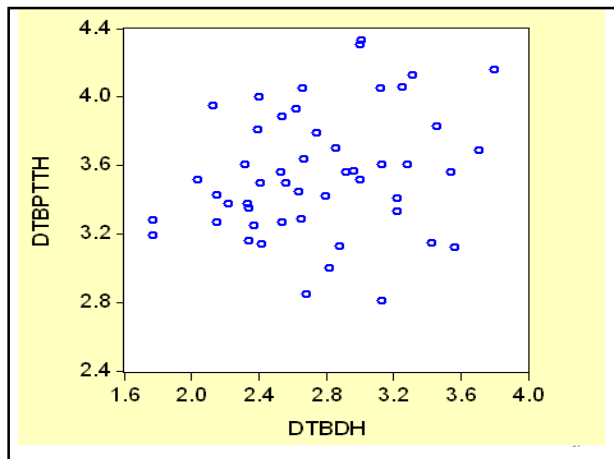
- Nếu $r_{X,Y} > 0$, nghĩa là $r_{X,Y} \approx 1$, thì X và Y có *quan hệ đồng biến chặt*.

- Ngược lại, nếu $r_{X,Y} < 0$, nghĩa là $r_{X,Y} \approx -1$, thì X và Y có *quan hệ nghịch biến chặt*.

15

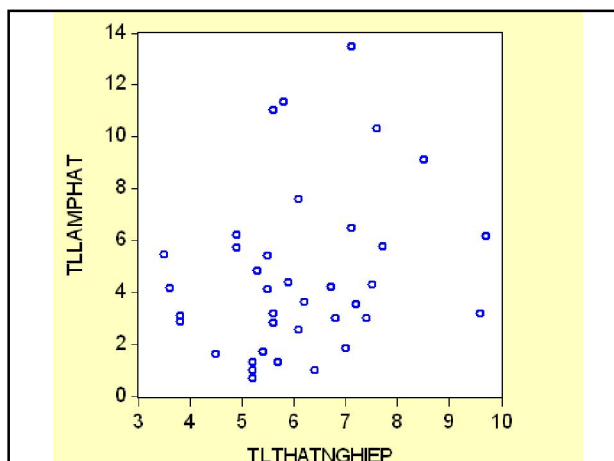
Ví dụ. Khảo sát mối quan hệ giữa điểm điểm trung bình ở PTTH và Đại học của 50 học sinh ở một trường đại học của Mỹ ta có bảng số liệu sau :

ĐH	PTTH	ĐH	PTTH	ĐH	PTTH	ĐH	PTTH	ĐH	PTTH
2.8	3.42	2.66	4.05	2.33	3.38	3	4.31	2.74	3.79
3.54	3.56	2.96	3.57	3.8	4.16	3.71	3.69	2.41	3.5
2.88	3.13	2.34	3.35	3.22	3.33	3.43	3.15	2.86	3.7
2.15	3.27	3.13	3.61	2.53	3.56	3.22	3.41	2.56	3.5
2.22	3.38	3.46	3.83	2.37	3.25	2.82	3	3.28	3.61
3.31	4.13	2.92	3.56	3.12	4.05	2.64	3.45	2.34	3.16
2.13	3.95	2.15	3.43	2.54	3.27	2.4	4	2.67	3.64
2.39	3.81	3	3.52	3.56	3.12	1.77	3.28	2.62	3.93
3.01	4.33	2.42	3.14	3.25	4.06	3.13	2.81	2.54	3.89
2.68	2.85	1.77	3.19	2.32	3.61	2.65	3.29	2.04	3.52



Ví dụ. Bảng sau cho số liệu về tỷ lệ lạm phát và tỷ lệ thất nghiệp của Mỹ từ năm 1959 đến 1995

Năm	TLLP	TLTN	Năm	TLLP	TLTN	Năm	TLLP	TLTN
1959	0.69	5.2	1971	4.38	5.9	1983	3.21	9.6
1960	1.72	5.4	1972	3.21	5.6	1984	4.32	7.5
1961	1.01	6.4	1973	6.22	4.9	1985	3.56	7.2
1962	1	5.2	1974	11.04	5.6	1986	1.86	7
1963	1.32	5.7	1975	9.13	8.5	1987	3.65	6.2
1964	1.31	5.2	1976	5.76	7.7	1988	4.14	5.5
1965	1.61	4.5	1977	6.5	7.1	1989	4.82	5.3
1966	2.86	3.8	1978	7.59	6.1	1990	5.4	5.5
1967	3.09	3.8	1979	11.35	5.8	1991	4.21	6.7
1968	4.19	3.6	1980	13.5	7.1	1992	3.01	7.4
1969	5.46	3.5	1981	10.32	7.6	1993	2.99	6.8
1970	5.72	4.9	1982	6.16	9.7	1994	2.56	6.1
						1995	2.83	5.6



4. Kiểm định mối tương quan tuyến tính của hai đại lượng X, Y.

Bài toán kiểm định

$$\begin{cases} H_0 : \rho_{X,Y} = 0 \\ H_1 : \rho_{X,Y} \neq 0 \end{cases}$$

❖ Nếu H_0 đúng, ta có thống kê

$$T = r_{x,y} \sqrt{\frac{n-2}{1-r_{x,y}^2}} \sim St(n-2)$$

4. Kiểm định mối tương quan tuyến tính của hai đại lượng X, Y.

Với mức ý nghĩa α cho trước ta có

$$C = t_{\frac{\alpha}{2}}^{n-2}$$

Nếu $|T| > C$ bác bỏ H_0 , nghĩa là X, Y có tương quan tuyến tính với nhau.

21

Ví dụ : Với số liệu ở ví dụ trên, ta có

$$r_{X,Y} = 0,2465; n = 37$$

Bài toán kiểm định

$$\begin{cases} H_0 : \rho_{X,Y} = 0 & \text{(LP và TN không có tương quan tuyến tính)} \\ H_1 : \rho_{X,Y} \neq 0 & \text{(LP và TN có tương quan tuyến tính)} \end{cases}$$

Nếu H_0 đúng, ta có thống kê

$$T = r \sqrt{\frac{n-2}{1-r^2}} \sim \text{St}(n-2), \quad T = 1,5047$$

Với $\alpha = 0,05$, ta tìm được : $C = t_{0,025}^{35} = 1,96$

Ta có $|T| > C$, bác bỏ H_0 .

22

Bài Giảng
KINH TẾ LƯỢNG
 (Econometric)

Chương 1
Hồi Quy Hai Biến

GV: ThS. Nguyễn Trung Đông
 nguyentrungdong144@yahoo.com

Chương 1. Hồi Quy Hai Biến

- ❖ Phân tích hồi quy
- ❖ Mô hình hồi quy
- ❖ Hệ số xác định mô hình
- ❖ Khoảng ước lượng
- ❖ Kiểm định sự phù hợp mô hình
- ❖ Bài toán dự báo

2

1. Phân tích hồi quy

- ❖ Nghiên cứu mối liên hệ phụ thuộc của một biến *phụ thuộc* (Y), theo một hay nhiều biến *độc lập* (X_i) khác.
- ❖ Phân tích hồi quy giải quyết các vấn đề sau
 - Ước lượng và dự đoán giá trị trung bình của biến phụ thuộc với giá trị đã cho của biến độc lập.
 - Kiểm định giả thuyết về bản chất của biến phụ thuộc.

3

1. Phân tích hồi quy

Chú ý:

- ❖ Biến độc lập là biến phi ngẫu nhiên.
- ❖ Biến phụ thuộc là biến ngẫu nhiên nó có phân phối xác định.
- Nghĩa là ứng với mỗi giá trị của biến độc lập, biến phụ thuộc có thể lấy giá trị khác nhau nhưng các giá trị này tuân theo luật phân phối xác định.

4

2. Mô Hình Hồi Quy

1. Hàm hồi quy tổng thể PRF

PRF=Population Regression Function

Ta xét PRF là hàm **tuyến tính** có dạng

$$E(Y | X = X_i) = \beta_1 + \beta_2 X_i \quad (1)$$

hay

$$E(Y | X) = \beta_1 + \beta_2 X$$

5

2. Mô Hình Hồi Quy

1. Hàm hồi quy tổng thể PRF

Hàm hồi quy tổng thể ngẫu nhiên của (1)

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

hay
$$Y = \beta_1 + \beta_2 X + \varepsilon$$

Trong đó $\beta_1, \beta_2, \varepsilon$ lần lượt là hệ số hồi quy và sai số ngẫu nhiên tổng thể.

6

2. Mô Hình Hồi Quy

2. Hàm hồi quy mẫu SRF

SRF=Sample Regression Function

Ta xét hàm hồi quy mẫu có dạng

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i \quad (2)$$

hay
$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X$$

Trong đó $\hat{Y}, \hat{\beta}_1, \hat{\beta}_2$ lần lượt là các ước lượng điểm của $E(Y|X), \beta_1, \beta_2$.

7

2. Mô Hình Hồi Quy

2. Hàm hồi quy mẫu SRF

Dạng ngẫu nhiên (2)

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + e_i$$

Với $e_i = Y_i - \hat{Y}_i$ là ước lượng điểm của ε_i (phần dư).

8

2. Mô Hình Hồi Quy

3. Tính chất của SRF

i) $\bar{Y} = \hat{\beta}_1 + \hat{\beta}_2 \bar{X}$; ii) $\bar{\hat{Y}} = \bar{Y}$

iii) $\bar{e} = \frac{1}{n} \sum_{i=1}^n e_i = 0$; iv) $\sum_{i=1}^n e_i \hat{Y}_i = 0$

Phần dư e và \hat{Y} không tương quan

v) $\sum_{i=1}^n e_i X_i = 0$

Phần dư e và X không tương quan

9

2. Mô Hình Hồi Quy

4. Phương pháp OLS

Giả sử $Y = \beta_1 + \beta_2 X$ là PRF cần tìm.

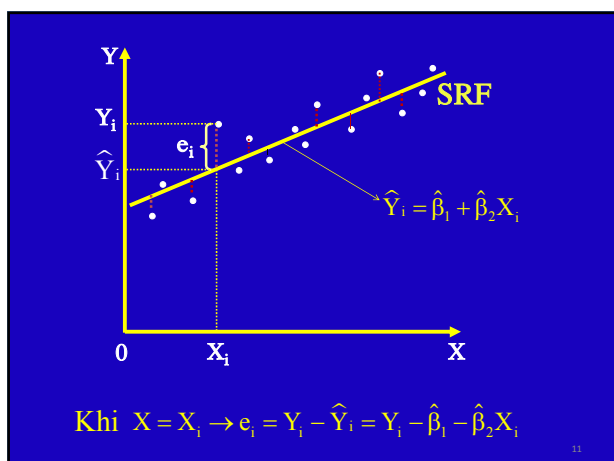
Ta ước lượng PRF bởi SRF có dạng

$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X$$

Từ một mẫu gồm n quan sát

(X_i, Y_i) ; $i=1, 2, \dots, n$, khi đó với mỗi i , ta có $e_i \equiv Y_i - \hat{Y}_i = Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i$ là các phần dư

10



11

2. Mô Hình Hồi Quy

Phương pháp OLS nhằm xác định các tham số $(\hat{\beta}_1, \hat{\beta}_2)$ sao cho :

$$f(\hat{\beta}_1, \hat{\beta}_2) \equiv \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2 \rightarrow \min$$

Khi đó $(\hat{\beta}_1, \hat{\beta}_2)$ thoả mãn hệ sau

$$\begin{cases} n\hat{\beta}_1 + \hat{\beta}_2 \sum_{i=1}^n X_i = \sum_{i=1}^n Y_i \\ \hat{\beta}_1 \sum_{i=1}^n X_i + \hat{\beta}_2 \sum_{i=1}^n X_i^2 = \sum_{i=1}^n X_i Y_i \end{cases}$$

12

2. Mô Hình Hồi Quy

Giải hệ trên ta được

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} = \frac{S_{X,Y}}{S_X^2} = r_{X,Y} \frac{S_Y}{S_X}$$

và $\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}$

13

2. Mô Hình Hồi Quy

Ví dụ 1. Bảng sau cho số liệu về lãi suất ngân hàng (Y) và tỷ lệ lạm phát (X) trong năm 1988 ở 9 nước.

X	7.2	4.0	3.1	1.6	4.8	51.0	2.0	6.6	4.4
Y	11.9	9.4	7.5	4.0	11.3	66.3	2.2	10.3	7.6

Với số liệu trên, ta tìm được (sử dụng MT)

$$\hat{\beta}_1 = 2.7417 \text{ và } \hat{\beta}_2 = 1.2494$$

Hay mô hình hồi quy : $\hat{Y} = 2.74 + 1.25 \cdot X^{14}$

Lập bảng ta tính được các tổng như sau

X_i	Y_i	$X_i Y_i$	X_i^2
7,2	11,9	85,68	51,84
4	9,4	37,6	16
3,1	7,5	23,25	9,61
1,6	4	6,4	2,56
4,8	11,3	54,24	23,04
51	66,3	3381,3	2601
2	2,2	4,4	4
6,6	10,3	67,98	43,56
4,4	7,6	33,44	19,36
$\sum X_i = 84,7$	$\sum Y_i = 130,5$	$\sum X_i Y_i = 3694,29$	$\sum X_i^2 = 2770,97$

$$\begin{cases} 9\hat{\beta}_1 + 84,7\hat{\beta}_2 = 130,5 \\ 84,7\hat{\beta}_1 + 2770,97\hat{\beta}_2 = 3694,29 \end{cases} \Leftrightarrow \begin{cases} \hat{\beta}_1 = 2,7417 \\ \hat{\beta}_2 = 1,2494 \end{cases}$$

15

2. Mô Hình Hồi Quy

5. Các giả thuyết của mô hình

GT1: Biến X là biến phi ngẫu nhiên.

GT2: $E(\varepsilon_i) = E(\varepsilon_i | X = X_i) = 0$.

GT3: $\text{Var}(\varepsilon_i) = \text{Var}(\varepsilon_j) = \sigma^2$, với mọi i, j

GT4: $\text{Cov}(\varepsilon_i, \varepsilon_j) = 0$

GT5: $\text{Cov}(\varepsilon_i, X_j) = 0$

GT6: $\varepsilon_i \sim N(0, \sigma^2)$

GT7: $Y_i \sim N(\beta_1 + \beta_2 X_i, \sigma^2)$

16

2. Mô Hình Hồi Quy

6. Tính chất các hệ số hồi quy

Các hệ số hồi quy có các tính chất sau:

- ❖ $\hat{\beta}_1$ và $\hat{\beta}_2$ được xác định một cách duy nhất ứng với các mẫu.
- ❖ $\hat{\beta}_1$ và $\hat{\beta}_2$ là các ước lượng điểm của β_1 và β_2 .
- ❖ Các hệ số hồi quy có phân phối sau:

17

2. Mô Hình Hồi Quy

$$\hat{\beta}_1 \sim N(\beta_1; \sigma_{\hat{\beta}_1}^2); \hat{\beta}_2 \sim N(\beta_2; \sigma_{\hat{\beta}_2}^2)$$

$$\text{và } Y = \frac{(n-2)\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-2)$$

Trong đó, các phương sai của các hệ số hồi quy được tính bởi các công thức sau :

18

2. Mô Hình Hồi Quy

$$\text{var}(\hat{\beta}_1) = \left[\frac{1}{n} + \frac{\bar{X}^2}{nS_X^2} \right] \sigma^2; \text{var}(\hat{\beta}_2) = \frac{\sigma^2}{nS_X^2}$$

Trong đó, σ^2 chưa biết ta thay σ^2 bởi ước lượng không chệch của nó là

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^n e_i^2 = \frac{n}{n-2} (1 - r_{X,Y}^2) S_Y^2$$

19

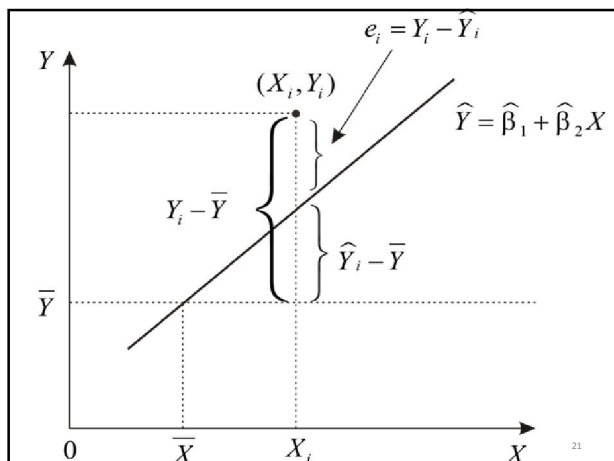
3. Hệ Số Xác Định Mô Hình

$$\text{TSS} = \sum_{i=1}^n (Y_i - \bar{Y})^2 = nS_Y^2.$$

$$\text{ESS} = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 = \hat{\beta}_2^2 \sum_{i=1}^n (X_i - \bar{X})^2 = n\hat{\beta}_2^2 S_X^2.$$

$$\begin{aligned} \text{RSS} &= \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \\ &= \text{TSS} - \text{ESS} = n(1 - r_{X,Y}^2) S_Y^2. \end{aligned}$$

20



3. Hệ Số Xác Định Mô Hình

Hệ số xác định MH (*coefficient of determination*)

$$R^2 = 1 - \text{RSS}/\text{TSS} = \text{ESS}/\text{TSS}, \text{ hay } R^2 = r_{X,Y}^2$$

để đo mức độ phù hợp của hàm hồi quy.

- Khi $R^2 = 1$, ta nói mô hình giải thích được toàn bộ sự thay đổi của các quan sát.
- Khi $R^2 = 0$, ta nói mô hình không giải thích được gì.

Khi đó ta còn có công thức sau :

$$\hat{\sigma}^2 = \frac{n}{n-2} (1 - r_{X,Y}^2) S_Y^2 = \frac{\text{RSS}}{n-2}$$

22

Chẳng hạn như trong ví dụ 1, ta có thể tính được các tham số sau :

$$\hat{\sigma}^2 = 2.975456987$$

$$\text{var}(\hat{\beta}_1) = 0.464118722$$

$$\text{var}(\hat{\beta}_2) = 0.001507439097$$

$$\text{TSS} = nS_Y^2 = 3102.04$$

$$\text{ESS} = n\hat{\beta}_2^2 S_X^2 = 3081.211809$$

$$\text{RSS} = n(1 - r_{X,Y}^2) S_Y^2 = 20.82819405$$

$$R^2 = \frac{\text{ESS}}{\text{TSS}} = 0.993285647$$

23

Nếu ta dùng phần mềm Eview, ta có kết quả sau

Dependent Variable: Y
Method: Least Squares
Date: 02/02/08 Time: 10:13
Sample: 1 9
Included observations: 9

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.741695	0.681263	4.024432	0.0050
X	1.249407	0.038826	32.17985	0.0000

R-squared	0.993286	Mean dependent var	14.50000
Adjusted R-squared	0.992326	S.D. dependent var	19.69150
S.E. of regression	1.724951	Akaike info criterion	4.121404
Sum squared resid	20.82819	Schwarz criterion	4.165232
Log likelihood	-16.54632	F-statistic	1035.543
Durbin-Watson stat	1.806737	Prob(F-statistic)	0.000000

24

4. Khoảng ước lượng cho các hệ số hồi quy tổng thể

Ta dùng các thống kê sau

$$T = \frac{\hat{\beta}_j - \beta_j}{\text{se}(\hat{\beta}_j)} \sim \text{St}(n-2); j = 1, 2$$

Với α cho trước ta tìm được : $C = t_{\frac{\alpha}{2}}^{n-2}$
Khoảng ước lượng cho β_j

$$\beta_j \in \left[\hat{\beta}_j - C \text{se}(\hat{\beta}_j); \hat{\beta}_j + C \text{se}(\hat{\beta}_j) \right], j = 1, 2$$

25

Ví dụ 2: Với số liệu ở ví dụ 1, ta có

$$n = 9; \hat{\beta}_1 = 2,7417; \text{se}(\hat{\beta}_1) = 0,6813$$

$$\hat{\beta}_2 = 1,2494; \text{se}(\hat{\beta}_2) = 0,0388$$

Với $\alpha = 0,05$, ta tìm được : $C = t_{0,025}^7 = 2,365$
Khoảng ước lượng cho các hệ số hồi quy

$$\beta_1 \in \left[\hat{\beta}_1 - C \text{se}(\hat{\beta}_1); \hat{\beta}_1 + C \text{se}(\hat{\beta}_1) \right] = [1,130; 4,353]$$

$$\beta_2 \in \left[\hat{\beta}_2 - C \text{se}(\hat{\beta}_2); \hat{\beta}_2 + C \text{se}(\hat{\beta}_2) \right] = [1,158; 1,341]$$

26

5. Khoảng ước lượng cho phương sai của sai số ngẫu nhiên tổng thể

Ta dùng thống kê sau

$$Y = \frac{(n-2)\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-2)$$

Với α ta có $a = \chi_{1-\frac{\alpha}{2}}^2(n-2)$; $b = \chi_{\frac{\alpha}{2}}^2(n-2)$

KUL cho σ^2 : $\sigma^2 \in \left[\frac{(n-2)\hat{\sigma}^2}{b}; \frac{(n-2)\hat{\sigma}^2}{a} \right]$

27

Ví dụ 3: Với số liệu ở ví dụ 1, ta có

$$n = 9; \hat{\sigma}^2 = 2,9755$$

Với $\alpha = 0,05$ ta có

$$a = \chi_{0,975}^2(7) = 1,69; b = \chi_{0,025}^2(7) = 16,013$$

KUL cho σ^2 :

$$\sigma^2 \in \left[\frac{(n-2)\hat{\sigma}^2}{b}; \frac{(n-2)\hat{\sigma}^2}{a} \right] = [1,301; 12,325]$$

28

6. Kiểm định sự phù hợp của mô hình

Bài toán kiểm định

$$\begin{cases} H_0 : \beta_2 = 0 \\ H_1 : \beta_2 \neq 0 \end{cases} \text{ hay } \begin{cases} H_0 : R^2 = 0 \\ H_1 : R^2 > 0 \end{cases}$$

Với H_0 khi biến độc lập X thay đổi không ảnh hưởng đến biến phụ thuộc Y (hay mô hình không phù hợp).

29

6. Kiểm định sự phù hợp của mô hình

Nếu H_0 đúng, ta có thống kê sau

$$T = \frac{\hat{\beta}_2}{\text{se}(\hat{\beta}_2)} \sim \text{St}(n-2) \quad \text{hay}$$

$$F = \frac{(n-2)R^2}{1-R^2} \sim F(1, n-2)$$

với α cho trước, ta có $C = t_{\frac{\alpha}{2}}^{n-2}$; $C = f_{\alpha}(1, n-2)$

Nếu $|T| > C$; $F > C$ bác bỏ H_0 .

30

Ví dụ 4: Với số liệu ở ví dụ 1, ta có

$$n = 9; \hat{\beta}_2 = 1,2494; \text{se}(\hat{\beta}_2) = 0,0388$$

Bài toán kiểm định

$$\begin{cases} H_0 : \beta_2 = 0 & \text{(LP thay đổi không ảnh hưởng tới LS)} \\ H_1 : \beta_2 \neq 0 & \text{(LP thay đổi ảnh hưởng tới LS)} \end{cases}$$

Nếu H_0 đúng, ta có thống kê

$$T = \frac{\hat{\beta}_2}{\text{se}(\hat{\beta}_2)} \sim \text{St}(n-2), \quad T = \frac{1,2494}{0,0388} = 32,201$$

Với $\alpha = 0,05$, ta tìm được: $C = t_{0,025}^7 = 2,365$

Ta có $|T| > C$, bác bỏ H_0 .

31

Ví dụ 5: Với số liệu ở ví dụ 1, ta có

$$n = 9; R^2 = 0,9933$$

Bài toán kiểm định

$$\begin{cases} H_0 : R^2 = 0 & \text{(Mô hình không phù hợp)} \\ H_1 : R^2 > 0 & \text{(Mô hình phù hợp)} \end{cases}$$

Ta dùng thống kê

$$F = \frac{(n-2)R^2}{1-R^2} \sim F(1, n-2), \quad F = 1036,91$$

Với $\alpha = 0,05$, ta tìm được: $C = f_{0,05}(1, 7) = 5,59$

Ta có $F > C$, bác bỏ H_0 .

32

6. Kiểm định sự phù hợp của mô hình

Phương pháp kiểm định bằng giá trị p - value được thực hiện như sau :

$$\text{Bước 1 : Tính giá trị } T_0 = \frac{\hat{\beta}_1 - \beta_1^*}{\text{se}(\hat{\beta}_1)}$$

Bước 2 : Tính p - value = $P(|T| > T_0)$, trong đó $T \sim \text{St}(n - 2)$

Bước 3 : So sánh giá trị p - value với mức ý nghĩa α cho trước, nếu

p - value $< \alpha$: Bác bỏ giả thuyết H.

33

6. Kiểm định sự phù hợp của mô hình

Tương tự cho bài toán kiểm định sự phù hợp của mô hình, với giá trị

$$F = \frac{(n - 2)R^2}{1 - R^2} \sim F(1, n - 2).$$

Thường thì các giá trị như

- p - value,
 - Các thống kê T, F
- được tính bằng phần mềm Eview.

34

7. Dự báo giá trị trung bình

Với $X = X_0$, ta có ước lượng điểm của Y

$$\hat{Y}_0 = \hat{\beta}_1 + \hat{\beta}_2 X_0$$

Để dự báo GTTB của Y, ta dùng thống kê

$$T = \frac{\hat{Y}_0 - E(Y | X = X_0)}{\text{se}(\hat{Y}_0)} \sim \text{St}(n - 2)$$

trong đó

$$\text{se}(\hat{Y}_0) = \sqrt{\text{var}(\hat{Y}_0)}$$

35

7. Dự báo giá trị trung bình

Với phương sai của \hat{Y}_0 được cho bởi

$$\text{var}(\hat{Y}_0) = \hat{\sigma}^2 \left[\frac{1}{n} + \frac{(X_0 - \bar{X})^2}{nS_x^2} \right]$$

Với α cho trước, ta có $C = t_{\frac{\alpha}{2}}^{n-2}$

Khoảng UL GTTB của Y:

$$E(Y | X = X_0) \in \left[\hat{Y}_0 - C \text{se}(\hat{Y}_0); \hat{Y}_0 + C \text{se}(\hat{Y}_0) \right]$$

36

Ví dụ 6: Với số liệu ở ví dụ 1, ta có

Với $X_0 = 5$, ta có $\hat{Y}_0 = 2,742 + 1,2494 \cdot 5 = 8,989$

Độ lệch chuẩn của \hat{Y}_0

$$se(\hat{Y}_0) = \sqrt{\text{var}(\hat{Y}_0)} = \sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{(X_0 - \bar{X})^2}{nS_x^2} \right]} = \sqrt{0,36} = 0,6$$

Với $\alpha = 0,05$, ta tìm được: $C = t_{0,025}^7 = 2,365$

Khoảng dự báo cho GTTB của Y

$$E(Y | X = 5) \in [\hat{Y}_0 - Cse(\hat{Y}_0); \hat{Y}_0 + Cse(\hat{Y}_0)] = [7,57; 10,41]$$

37

8. Dự báo giá trị cá biệt Y_0

Để báo cho giá trị cá biệt Y_0 , ta dùng thống kê sau

$$T = \frac{Y_0 - \hat{Y}_0}{se(Y_0 - \hat{Y}_0)} \sim \text{St}(n - 2)$$

Trong đó

$$se(Y_0 - \hat{Y}_0) = \sqrt{\text{var}(Y_0 - \hat{Y}_0)}$$

38

8. Dự báo giá trị cá biệt Y_0

Với phương sai của $(Y_0 - \hat{Y}_0)$ được cho bởi

$$\text{var}(Y_0 - \hat{Y}_0) = \hat{\sigma}^2 + \text{var}(\hat{Y}_0)$$

Với α cho trước, ta có $C = t_{\frac{\alpha}{2}}^{n-2}$

Khoảng UL GTCB của Y:

$$Y_0 \in [\hat{Y}_0 - Cse(Y_0 - \hat{Y}_0); \hat{Y}_0 + Cse(Y_0 - \hat{Y}_0)]$$

39

Ví dụ 7: Với số liệu ở ví dụ 1, ta có

Với $X_0 = 5$, ta có $\hat{Y}_0 = 2,742 + 1,2494 \cdot 5 = 8,989$

Độ lệch chuẩn của $(Y_0 - \hat{Y}_0)$

$$se(Y_0 - \hat{Y}_0) = \sqrt{\text{var}(Y_0 - \hat{Y}_0)} = \sqrt{\hat{\sigma}^2 + \text{var}(\hat{Y}_0)} = 1,83$$

Với $\alpha = 0,05$, ta tìm được: $C = t_{0,025}^7 = 2,365$

Khoảng dự báo cho GTCB của Y

$$Y_0 \in [\hat{Y}_0 - Cse(Y_0 - \hat{Y}_0); \hat{Y}_0 + Cse(Y_0 - \hat{Y}_0)] = [4,66; 13,32]$$

40

Ví dụ 8: Cho số liệu về năng suất (Y: tạ/ha) và mức phân bón (X: tạ/ha) của một loại cây trồng từ năm 1988 đến năm 1997 như sau.

Năm	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
X	6	10	12	14	16	18	22	24	26	32
Y	40	44	46	48	52	58	60	68	74	80

Giả sử X và Y có quan hệ tuyến tính

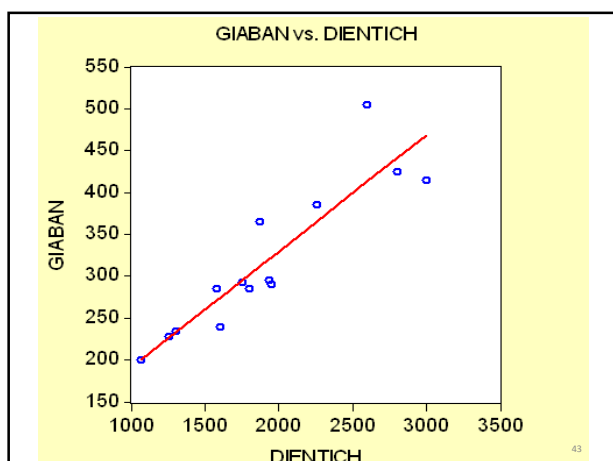
41

Ví dụ 9: Bảng sau cho số liệu về giá bán một căn nhà (Y: ngàn USD/ft²) và diện tích (X: ft²) như sau:

Diện tích	Giá bán	Diện tích	Giá bán
1065	199.9	1870	365
1254	228	1935	295
1300	235	1948	290
1577	285	2254	385
1600	239	2600	505
1750	293	2800	425
1800	285	3000	415

Giả sử X và Y có quan hệ tuyến tính

42



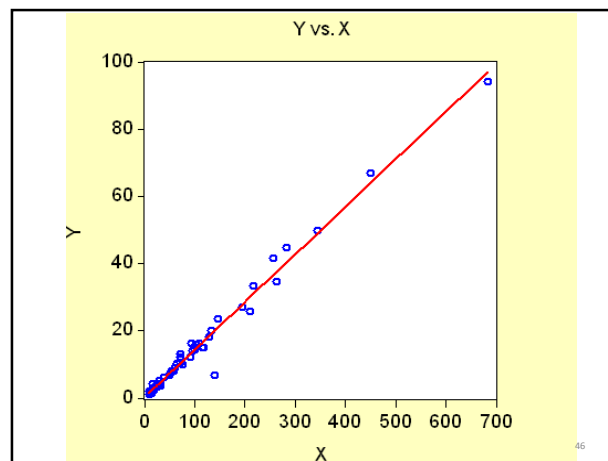
Dependent Variable: GIABAN
Method: Least Squares
Date: 03/07/08 Time: 14:13
Sample: 1 14
Included observations: 14

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	52.35091	37.28549	1.404056	0.1857
DIENTICH	0.138750	0.018733	7.406788	0.0000
R-squared	0.820522	Mean dependent var	317.4929	
Adjusted R-squared	0.805565	S.D. dependent var	88.49816	
S.E. of regression	39.02304	Akaike info criterion	10.29774	
Sum squared resid	18273.57	Schwarz criterion	10.38904	
Log likelihood	-70.08421	F-statistic	54.86051	
Durbin-Watson stat	1.975057	Prob(F-statistic)	0.000008	

44

Ví dụ 10: Cho số liệu về thu nhập (X: ngàn USD/tháng) và chi tiêu cho việc chăm sóc sức khỏe (Y: ngàn USD/tháng) của 51 cá nhân ở Mỹ. Ta có bảng kết quả xuất ra từ Eview như sau (slide kế tiếp)
Giả sử X và Y có tương quan tuyến tính với nhau. Dựa vào bảng kết quả trả lời các câu hỏi sau

45



46

Dependent Variable: Y
Method: Least Squares
Date: 02/02/08 Time: 11:21
Sample: 1 51
Included observations: 51

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.176496	0.467509	0.377525	0.7074
X	0.141652	0.002875	49.27179	0.0000
R-squared	0.980216	Mean dependent var	15.06886	
Adjusted R-squared	0.979812	S.D. dependent var	17.92664	
S.E. of regression	2.547102	Akaike info criterion	4.746215	
Sum squared resid	317.8986	Schwarz criterion	4.821973	
Log likelihood	-119.0285	F-statistic	2427.709	
Durbin-Watson stat	2.209485	Prob(F-statistic)	0.000000	

47

Với mức ý nghĩa 5%, hãy trả lời các câu hỏi.

- 1) Viết hàm SRF. Nêu ý nghĩa hệ số góc.
- 2) Tìm KTC cho các hệ số HQ tổng thể.
- 3) Hãy cho biết thu nhập thay đổi có ảnh hưởng đến chi tiêu cho sức khỏe không.
- 4) Giải thích ý nghĩa hệ số xác định mô hình và kiểm định sự phù hợp của mô hình.
- 5) Với mức thu nhập 100 nghìn USD. Hãy dự báo GTTB và GTCB của chi tiêu cho sức khỏe.

48

Bài Giảng
KINH TẾ LƯỢNG
 (Econometric)

Chương 2
Hồi Quy Bội
(nhiều biến)

GV: ThS. Nguyễn Trung Đông
 nguyentrungdong144@yahoo.com

Chương 2. Hồi Quy Bội

- ❖ Hàm hồi quy tổng thể PRF.
- ❖ Các giả mô hình thuyết.
- ❖ Ước lượng tham số.
- ❖ Hệ số xác định mô hình hồi quy bội.
- ❖ Ma trận tương quan, Ma trận hiệp phương sai.
- ❖ Khoảng tin cậy và kiểm định giả thuyết.
- ❖ Dự báo.
- ❖ Một số dạng hàm hồi quy.
- ❖ Hồi quy với biến giả.

1. Hàm hồi quy tổng thể PRF

Xét hàm hồi quy tuyến tính k biến

$$E(Y|X_2, X_3, \dots, X_k) = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k$$

Hay $Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k + \varepsilon$

Trong đó

ε là sai số ngẫu nhiên

β_1 là hệ số tự do

$\beta_2, \beta_3, \dots, \beta_k$ là các hệ số hồi quy riêng

1. Hàm hồi quy tổng thể PRF

Từ một mẫu quan sát $(Y_i, X_{2,i}, X_{3,i}, \dots, X_{k,i})$

với $i = 1, 2, \dots, n$, lấy từ tổng thể, ta có hệ sau

$$\begin{cases} Y_1 = \beta_1 + \beta_2 X_{2,1} + \dots + \beta_k X_{k,1} + e_1 \\ Y_2 = \beta_1 + \beta_2 X_{2,2} + \dots + \beta_k X_{k,2} + e_2 \\ \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \\ Y_n = \beta_1 + \beta_2 X_{2,n} + \dots + \beta_k X_{k,n} + e_n \end{cases}$$

Với e_j là các phần dư của số hạng thứ j .

1. Hàm hồi quy tổng thể PRF

Viết hệ trên dưới dạng ma trận như sau

$$Y = X \cdot \beta + e$$

Trong đó

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_n \end{pmatrix}; \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_k \end{pmatrix}; e = \begin{pmatrix} e_1 \\ e_2 \\ \dots \\ e_n \end{pmatrix}$$

$$X = \begin{pmatrix} 1 & X_{2,1} & X_{3,1} & \dots & X_{k,1} \\ 1 & X_{2,2} & X_{3,2} & \dots & X_{k,2} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & X_{2,n} & X_{3,n} & \dots & X_{k,n} \end{pmatrix}$$

5

2. Các giả thuyết mô hình

$$GT1 : E(e_i) = 0, \forall i$$

$$GT2 : E(e_i, e_j) = \begin{cases} 0 & \text{khi } i \neq j \\ \sigma^2 & \text{khi } i = j \end{cases}$$

Hay dưới dạng ma trận $E(ee^T) = \sigma^2 I$

GT3 : Các biến độc lập phi ngẫu nhiên.

GT4 : Không có hiện tượng cộng tuyến giữa các biến độc lập.

6

3. Ước lượng tham số

Xét hàm hồi quy mẫu SRF có dạng

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2,i} + \hat{\beta}_3 X_{3,i} + \dots + \hat{\beta}_k X_{k,i} + e_i$$

Hay dưới dạng ma trận $Y = X\hat{\beta} + e$
trong đó

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \dots \\ \hat{\beta}_k \end{pmatrix}; e = \begin{pmatrix} e_1 \\ e_2 \\ \dots \\ e_k \end{pmatrix} = Y - X\hat{\beta}$$

7

3. Ước lượng tham số

Khi đó, phương pháp OLS, xác định các hệ số hồi quy sao cho

$$RSS = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

$$RSS = \sum_{i=1}^n (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_{2,i} - \dots - \hat{\beta}_k X_{k,i})^2 \rightarrow \min$$

8

3. Ước lượng tham số

Khi đó các tham số hồi quy thỏa mãn hệ

$$\frac{\partial \text{RSS}}{\partial \hat{\beta}} = 0 \Leftrightarrow (X^T X) \hat{\beta} = X^T Y \Leftrightarrow \hat{\beta} = (X^T X)^{-1} (X^T Y)$$

Trong đó

$$X^T \cdot X = \begin{pmatrix} n & \sum_{i=1}^n X_{2,i} & \dots & \sum_{i=1}^n X_{k,i} \\ \sum_{i=1}^n X_{2,i} & \sum_{i=1}^n X_{2,i}^2 & \dots & \sum_{i=1}^n X_{2,i} X_{k,i} \\ \dots & \dots & \dots & \dots \\ \sum_{i=1}^n X_{k,i} & \sum_{i=1}^n X_{k,i} X_{2,i} & \dots & \sum_{i=1}^n X_{k,i}^2 \end{pmatrix}; X^T Y = \begin{pmatrix} \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_{2,i} Y_i \\ \dots \\ \sum_{i=1}^n X_{k,i} Y_i \end{pmatrix}$$

Ví dụ 1. Số liệu quan sát của một mẫu cho ở bảng sau. Trong đó

Y: Lượng hàng bán được của một loại hàng hóa (tấn / tháng)

X_2 : Thu nhập của người tiêu dùng (triệu / năm)

X_3 : Giá bán của loại hàng này (ngàn đ / kg)

Y	X_2	X_3	Y	X_2	X_3
20	8	2	17	6	5
18	7	3	16	5	6
19	8	4	15	5	7
18	8	4	13	4	8
17	6	5	12	3	8

10

Từ bảng số liệu trên ta tính được các tổng

$$\sum Y_i = 165; \sum X_{2i} = 60; \sum X_{3i} = 52; \sum Y_i^2 = 2781; \sum X_{2i}^2 = 388;$$

$$\sum X_{3i}^2 = 308; \sum X_{2i} X_{3i} = 282; \sum Y_i X_{2i} = 1029; \sum Y_i X_{3i} = 813;$$

$$(X^T X)^{-1} = \begin{pmatrix} 10 & 60 & 52 \\ 60 & 388 & 282 \\ 52 & 282 & 308 \end{pmatrix}^{-1} = \frac{1}{1528} \begin{pmatrix} 39980 & -3816 & -3256 \\ -3816 & 376 & 300 \\ -3256 & 300 & 280 \end{pmatrix}$$

$$\hat{\beta} = \frac{1}{1528} \begin{pmatrix} 39980 & -3816 & -3256 \\ -3816 & 376 & 300 \\ -3256 & 300 & 280 \end{pmatrix} \begin{pmatrix} 165 \\ 1029 \\ 813 \end{pmatrix} = \begin{pmatrix} 22908 / 1528 \\ 1164 / 1528 \\ -900 / 1528 \end{pmatrix}$$

Hay
$$\hat{\beta} = \begin{pmatrix} 14.99215 \\ 0.76178 \\ -0.58901 \end{pmatrix}$$

Vậy hàm hồi quy cần tìm là $\hat{Y}_i = 14.99215 + 0.76178X_{2i} - 0.58901X_{3i}$

Kết quả tính toán trên cho bởi phần mềm Eview

Dependent Variable: LUONGHANG

Method: Least Squares

Date: 04/04/08 Time: 14:19

Sample: 1 10

Included observations: 10

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	14.99215	2.922713	5.129531	0.0014
THUNHAP	0.761780	0.283438	2.687639	0.0312
GIABAN	-0.589005	0.244593	-2.408105	0.0469

R-squared	0.960934	Mean dependent var	16.50000
Adjusted R-squared	0.949773	S.D. dependent var	2.549510
S.E. of regression	0.571382	Akaike info criterion	1.961807
Sum squared resid	2.285340	Schwarz criterion	2.052582
Log likelihood	-6.809035	F-statistic	86.09278
Durbin-Watson stat	1.800073	Prob(F-statistic)	0.000012

4. Hệ số xác định MH hồi quy bội

Để đánh giá mức độ phù hợp của mô hình hồi quy, ta dùng hệ số xác định R^2 được xác định như sau

$$R^2 = 1 - \frac{RSS}{TSS} = \frac{ESS}{TSS}$$

Trong đó

$$TSS = Y^T Y - n(\bar{Y})^2 = nS_Y^2$$

$$ESS = \hat{\beta}^T X^T Y - n(\bar{Y})^2$$

$$RSS = TSS - ESS.$$

13

4. Hệ số xác định MH hồi quy bội

❖ Ý nghĩa của R^2 cũng tương tự như trong mô hình hai biến.

❖ Để so sánh mức độ phù hợp của các mô hình có số biến độc lập khác nhau, hay

❖ Để xem xét việc có nên đưa thêm các biến độc lập mới vào mô hình không.

❖ Khi đó ta dùng hệ số xác định điều chỉnh là: $\bar{R}^2 = 1 - (1 - R^2) \frac{n-1}{n-k}$

Biến độc lập đưa vào mô hình là có ý nghĩa nếu làm tăng giá trị của \bar{R}^2 .

14

5. Ma trận tương quan

$$R = \begin{pmatrix} 1 & r_{1,2} & \dots & r_{1,k} \\ r_{2,1} & 1 & \dots & r_{2,k} \\ \dots & \dots & \dots & \dots \\ r_{k,1} & r_{k,2} & \dots & 1 \end{pmatrix}$$

Trong đó

$$r_{i,j} = \frac{\sum_{i=1}^n y_i x_{j,i}}{\sqrt{\sum_{i=1}^n y_i^2 \sum_{i=1}^n x_{j,i}^2}}, \quad r_{i,j} = \frac{\sum_{i=1}^n x_{t,i} x_{j,i}}{\sqrt{\sum_{i=1}^n x_{t,i}^2 \sum_{i=1}^n x_{j,i}^2}}; \quad x_{j,i} = X_{j,i} - \bar{X}_j$$

15

6. Ma trận hiệp phương sai

$$\text{cov}(\hat{\beta}) = \begin{pmatrix} \text{var}(\hat{\beta}_1) & \text{cov}(\hat{\beta}_1, \hat{\beta}_2) & \dots & \text{cov}(\hat{\beta}_1, \hat{\beta}_k) \\ \text{cov}(\hat{\beta}_2, \hat{\beta}_1) & \text{var}(\hat{\beta}_2) & \dots & \text{cov}(\hat{\beta}_2, \hat{\beta}_k) \\ \dots & \dots & \dots & \dots \\ \text{cov}(\hat{\beta}_k, \hat{\beta}_1) & \text{cov}(\hat{\beta}_k, \hat{\beta}_2) & \dots & \text{var}(\hat{\beta}_k) \end{pmatrix}$$

Ta tính $\text{cov}(\hat{\beta}) = \sigma^2 (X^T \cdot X)^{-1}$ ta thay σ^2

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

16

Ví dụ 2. với số liệu cho trong ví dụ 1, ta có

$$TSS = Y^T Y - n(\bar{Y})^2 = 58.5$$

$$ESS = \hat{\beta}^T (X^T Y) - n(\bar{Y})^2 = 2778.71 - 10(16.5)^2 = 56.211$$

$$RSS = TSS - ESS = 58.5 - 56.21 = 2.289$$

$$R^2 = \frac{ESS}{TSS} = \frac{56.211}{58.5} = 0.96087$$

$$\hat{\sigma}^2 = \frac{RSS}{n-3} = \frac{2.289}{7} = 0.327$$

17

6. Ma trận hiệp phương sai

Vậy, ta có ma trận hiệp phương sai

$$\begin{aligned} \text{cov}(\hat{\beta}) &= \hat{\sigma}^2 (X^T \cdot X)^{-1} \\ &= \frac{0.327}{1528} \begin{pmatrix} 39980 & -3816 & -3256 \\ -3816 & 376 & 300 \\ -3256 & 300 & 280 \end{pmatrix} \\ &= \begin{pmatrix} 8.55593 & -0.81664 & -0.6968 \\ -0.81664 & 0.080466 & 0.0642 \\ -0.6968 & 0.0642 & 0.05992 \end{pmatrix} \end{aligned}$$

18

Các kết quả tính ở trên được cho bởi Eview như

Mô hình hồi quy

Estimation Command: LS Y C X2 X3

Estimation Equation: Y = C(1) + C(2)*X2 + C(3)*X3

Substituted Coefficients: Y = 14.9921466 + 0.7617801047*X2 - 0.5890052356*X3

Coefficient Covariance Matrix			
	C	X2	X3
C	8.542250	-0.815338	-0.695687
X2	-0.815338	0.080337	0.064099
X3	-0.695687	0.064099	0.059826

Correlation Matrix			
	X2	X3	Y
X2	1.000000	-0.924588	0.963624
X3	-0.924588	1.000000	-0.959490
Y	0.963624	-0.959490	1.000000

19

7. Khoảng tin cậy cho các hệ số hồi quy tổng thể

Ta dùng thống kê sau

$$T = \frac{\hat{\beta}_j - \beta_j}{\text{se}(\hat{\beta}_j)} \sim \text{St}(n - k)$$

Trong đó $\text{se}(\hat{\beta}_j) = \sqrt{\text{var}(\hat{\beta}_j)}$

Được cho trong ma trận hiệp phương sai

20

7. Khoảng tin cậy cho các hệ số hồi quy tổng thể

Với mức ý nghĩa α cho trước, ta có

$$C = t_{\frac{\alpha}{2}}^{n-k}$$

Khoảng ước lượng cho β_j , $j = 1, 2, \dots, k$

$$\beta_j \in \left[\hat{\beta}_j - C \text{se}(\hat{\beta}_j); \hat{\beta}_j + C \text{se}(\hat{\beta}_j) \right]$$

21

Ví dụ 3: Với số liệu ở ví dụ 1, ta có

$$n = 10; \hat{\beta}_1 = 14,992; \text{se}(\hat{\beta}_1) = 2,923; \hat{\beta}_2 = 0,762$$

$$\text{se}(\hat{\beta}_2) = 0,283; \hat{\beta}_3 = -0,589; \text{se}(\hat{\beta}_3) = 0,245$$

Với $\alpha = 0,05$, ta tìm được: $C = t_{0,025}^7 = 2,365$

Khoảng ước lượng cho các hệ số hồi quy

$$\beta_1 \in \left[\hat{\beta}_1 - C \text{se}(\hat{\beta}_1); \hat{\beta}_1 + C \text{se}(\hat{\beta}_1) \right] = [8,0791; 21,9049]$$

$$\beta_2 \in \left[\hat{\beta}_2 - C \text{se}(\hat{\beta}_2); \hat{\beta}_2 + C \text{se}(\hat{\beta}_2) \right] = [0,0927; 1,4313]$$

$$\beta_3 \in \left[\hat{\beta}_3 - C \text{se}(\hat{\beta}_3); \hat{\beta}_3 + C \text{se}(\hat{\beta}_3) \right] = [-1,1684; -0,0096]$$

22

8. Khoảng ước lượng cho phương sai của sai số ngẫu nhiên tổng thể

Ta dùng thống kê sau

$$Y = \frac{(n-k)\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-k)$$

Với α ta có $a = \chi_{1-\frac{\alpha}{2}}^2(n-k)$; $b = \chi_{\frac{\alpha}{2}}^2(n-k)$

KUL cho σ^2 :
$$\sigma^2 \in \left[\frac{(n-k)\hat{\sigma}^2}{b}; \frac{(n-k)\hat{\sigma}^2}{a} \right]$$

23

Ví dụ 4: Với số liệu ở ví dụ 1, ta có

$$n = 10; k = 3; \hat{\sigma}^2 = (0,571382)^2 = 0,3265$$

Với $\alpha = 0,05$ ta có

$$a = \chi_{0,975}^2(7) = 1,69; b = \chi_{0,025}^2(7) = 16,013$$

KUL cho σ^2 :

$$\sigma^2 \in \left[\frac{(n-k)\hat{\sigma}^2}{b}; \frac{(n-k)\hat{\sigma}^2}{a} \right] = [0,143; 1,352]$$

24

9. Kiểm định sự phù hợp của mô hình

Kiểm định giả thuyết (KĐ toàn phần)

$$H_0 : \beta_2 = \beta_3 = \dots = \beta_k = 0 \Leftrightarrow H_0 : R^2 = 0$$

Ta dùng thống kê sau :

$$F = \frac{\frac{ESS}{k-1}}{\frac{RSS}{n-k}} = \frac{\frac{R^2}{k-1}}{\frac{1-R^2}{n-k}} \sim F(k-1; n-k)$$

Với α cho trước, ta có : $C = f_{\alpha}(k-1; n-k)$

Nếu $F > C$: bác bỏ H_0 .

25

Ví dụ 5: Với số liệu ở ví dụ 1, ta có

$$n = 10; k = 3; R^2 = 0,96087$$

Bài toán kiểm định

$$\begin{cases} H_0 : R^2 = 0 & \text{(Mô hình không phù hợp)} \\ H_1 : R^2 > 0 & \text{(Mô hình phù hợp)} \end{cases}$$

Ta dùng thống kê

$$F = \frac{(n-k)R^2}{(k-1)(1-R^2)} \sim F(k-1, n-k), \quad F = 86,093$$

Với $\alpha = 0,05$, ta tìm được: $C = f_{0,05}(2, 7) = 4,74$

Ta có $F > C$, bác bỏ H_0 .

26

9. Kiểm định sự phù hợp của mô hình

Kiểm định giả thuyết (KĐ từng phần)

$$H_0 : \beta_j = 0; \quad j = 2, 3, \dots, k$$

Nếu H_0 đúng, ta có thống kê sau :

$$T = \frac{\hat{\beta}_j}{\text{Se}(\hat{\beta}_j)} \sim \text{St}(n-k)$$

Với α cho trước, ta có : $C = t_{\frac{\alpha}{2}}^{n-k}$

Nếu $|T| > C$: bác bỏ H_0 .

27

Ví dụ 6: Với số liệu ở ví dụ 1, ta có

$$n = 10; \hat{\beta}_3 = -0,589; \text{se}(\hat{\beta}_3) = 0,245$$

Bài toán kiểm định

$$\begin{cases} H_0 : \beta_3 = 0 & \text{(Giá bán thay đổi không ảnh hưởng tới lượng hàng)} \\ H_1 : \beta_3 \neq 0 & \text{(Giá bán thay đổi làm ảnh hưởng tới lượng hàng)} \end{cases}$$

Nếu H_0 đúng, ta có thống kê

$$T = \frac{\hat{\beta}_3}{\text{se}(\hat{\beta}_3)} \sim \text{St}(n-3), \quad T = -2,4041$$

Với $\alpha = 0,05$, ta tìm được : $C = t_{0,025}^7 = 2,365$

Ta có $|T| > C$, bác bỏ H_0 .

28

10. Dự báo

Dự báo cho giá trị trung bình

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

Với dự báo điểm là $\hat{Y}_0 = \hat{\beta}_1 + \hat{\beta}_2 X_2^0 + \dots + \hat{\beta}_k X_k^0$

Ta dùng thống kê sau

$$T = \frac{\hat{Y}_0 - E(Y|X=X^0)}{\text{se}(\hat{Y}_0)} \sim \text{St}(n-k)$$

Trong đó $\text{se}(\hat{Y}_0) = \sqrt{\text{var}(\hat{Y}_0)}$

29

10. Dự báo

Với phương sai của \hat{Y}_0

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

Với α cho trước, ta có $C = t_{\frac{\alpha}{2}}^{n-k}$

Khoảng ước lượng GTTB của Y

$$E(Y|X=X^0) \in [\hat{Y}_0 - C \text{se}(\hat{Y}_0); \hat{Y}_0 + C \text{se}(\hat{Y}_0)]$$

30

10. Dự báo

- Dự báo cho giá trị cá biệt Y_0

Ta dùng thống kê

$$T = \frac{Y_0 - \hat{Y}_0}{\text{se}(Y_0 - \hat{Y}_0)} \sim \text{St}(n-k)$$

Trong đó

$$\text{se}(Y_0 - \hat{Y}_0) = \sqrt{\text{var}(Y_0 - \hat{Y}_0)}$$

31

10. Dự báo

Với phương sai của $(Y_0 - \hat{Y}_0)$

$$\text{Var}(Y_0 - \hat{Y}_0) = \hat{\sigma}^2 + \text{Var}(\hat{Y}_0)$$

Với α cho trước, ta có $C = t_{\frac{\alpha}{2}}^{n-k}$

Khoảng ước lượng GTCB của Y

$$Y_0 \in [\hat{Y}_0 - C \text{se}(Y_0 - \hat{Y}_0); \hat{Y}_0 + C \text{se}(Y_0 - \hat{Y}_0)]$$

32

Ví dụ 7. Cho biết số liệu về sản lượng Y, phân hóa học X₂, thuốc trừ sâu X₃, tính trên một đơn vị diện tích ha, cho trong bảng sau

Y	40	44	46	48	52	58	60	68	74	80
X ₂	6	10	12	14	16	18	22	24	26	32
X ₃	4	4	5	7	9	12	14	20	21	24

33

Dependent Variable: SANLUONG
Method: Least Squares
Date: 04/04/08 Time: 19:47
Sample: 1 10
Included observations: 10

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	31.98067	1.631796	19.59845	0.0000
PHANBON	0.650051	0.250161	2.598527	0.0355
THUOC	1.109868	0.267434	4.150068	0.0043
R-squared	0.991634	Mean dependent var	57.00000	
Adjusted R-squared	0.989243	S.D. dependent var	13.47426	
S.E. of regression	1.397467	Akaike info criterion	3.750525	
Sum squared resid	13.67040	Schwarz criterion	3.841300	
Log likelihood	-15.75262	F-statistic	414.8492	
Durbin-Watson stat	2.114085	Prob(F-statistic)	0.000000	

34

11. Một số dạng hàm hồi quy

❖ Hàm sản xuất Cobb – Douglas (tuyến tính Log)

$$\text{Dạng tổng quát : } Y = \beta_1 X_2^{\beta_2} X_3^{\beta_3} \dots X_m^{\beta_m} e^\varepsilon$$

$$\text{Dạng thường dùng : } Y = \beta_1 X_2^{\beta_2} X_3^{\beta_3} e^\varepsilon$$

❖ Mô hình nghịch đảo

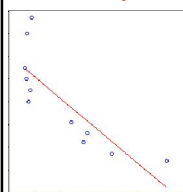
$$Y = \frac{1}{\beta_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon}$$

❖ Mô hình đa thức

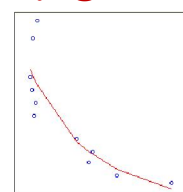
$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_2^2 + \dots + \beta_k X_2^k + \varepsilon$$

35

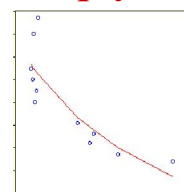
11. Một số dạng hàm hồi quy



Mô hình TT



Mô hình Nghịch



Mô hình Logarit

Variable	Coefficient	Variable	Coefficient	Variable	Coefficient
C	2.691124	C	1.577738	C	0.777418
X	-0.479529	1/X	0.578952	LOG(X)	-0.253046
R-squared	0.652757	R-squared	0.730893	R-squared	0.744800

12. Hồi quy với biến giả

➤ **Ví dụ 8.** Ta cần đánh giá sự khác biệt về mức tiền lương (Y), của các nhân viên, phụ thuộc vào giới tính. Khi đó, ta cần đưa vào mô hình hồi quy một biến giả D , với $D = 0$: Nữ và $D = 1$: Nam.

➤ (Lưu ý : nếu như ta cần so sánh n phạm trù khác nhau, ta cần có $n - 1$ biến giả)

Xét mô hình $Y_i = E(Y|D_i) = \beta_1 + \beta_2 D_i + \varepsilon_i$

với $D_i = 0 \Leftrightarrow E(Y|D_i = 0) = \beta_1$

$D_i = 1 \Leftrightarrow E(Y|D_i = 1) = \beta_1 + \beta_2$

37

So sánh hai hàm hồi quy

Giả sử, ta có hai bộ số liệu $(X_i, Y_i), i = \overline{1, n_1}$ và $(X_j, Y_j), j = \overline{1, n_2}$, ta sẽ có hai mô hình

$$Y_i = \lambda_1 + \lambda_2 X_i + \varepsilon_{1,i}, \quad i = \overline{1, n_1} \quad (1a)$$

$$Y_j = \gamma_1 + \gamma_2 X_j + \varepsilon_{2,j}, \quad j = \overline{1, n_2} \quad (1b)$$

Để kiểm định cho sự khác nhau của hai mô hình, ta dùng phép kiểm định Chow, như sau

38

Các bước kiểm định Chow

Bước 1: Tìm hàm hồi quy với mẫu

$n = n_1 + n_2$. Khi đó ta thu được RSS

Bước 2: Tìm hàm hồi quy riêng với mẫu n_1, n_2 . Tương tự ta cũng có RSS_1 và RSS_2

$$\overline{RSS} = RSS_1 + RSS_2$$

Bước 3: Ta dùng thống kê sau

$$F = \frac{(RSS - \overline{RSS}) / k}{\overline{RSS} / (n_1 + n_2 - 2k)} \sim F(k, n_1 + n_2 - 2k)$$

39

Ví dụ.: Khảo sát số tủ lạnh bán được tại Mỹ từ quý 1 năm 1978 đến quý 4 năm 1985 được cho trong bảng sau.

Năm: quý	FRIG	DUM1	DUM2	DUM3	Năm: quý	FRIG	DUM1	DUM2	DUM3
1978 :1	1317	1	0	0	1982 :1	943	1	0	0
1978 :2	1615	0	1	0	1982 :2	1175	0	1	0
1978 :3	1662	0	0	1	1982 :3	1269	0	0	1
1978 :4	1295	0	0	0	1982 :4	973	0	0	0
1979 :1	1271	1	0	0	1983 :1	1102	1	0	0
1979 :2	1555	0	1	0	1983 :2	1344	0	1	0
1979 :3	1639	0	0	1	1983 :3	1641	0	0	1
1979 :4	1238	0	0	0	1983 :4	1225	0	0	0
1980 :1	1277	1	0	0	1984 :1	1429	1	0	0
1980 :2	1258	0	1	0	1984 :2	1699	0	1	0
1980 :3	1417	0	0	1	1984 :3	1749	0	0	1
1980 :4	1185	0	0	0	1984 :4	1117	0	0	0
1981 :1	1196	1	0	0	1985 :1	1242	1	0	0
1981 :2	1410	0	1	0	1985 :2	1684	0	1	0
1981 :3	1417	0	0	1	1985 :3	1764	0	0	1
1981 :4	919	0	0	0	1985 :4	1328	0	0	0

Trong đó.: FRIG (ngàn cái)

DUM1 = $\begin{cases} 1 : \text{quý 1} \\ 0 : \text{quý khác} \end{cases}$; DUM2 = $\begin{cases} 1 : \text{quý 2} \\ 0 : \text{quý khác} \end{cases}$

DUM3 = $\begin{cases} 1 : \text{quý 3} \\ 0 : \text{quý khác} \end{cases}$

40

Dependent Variable: FRIG				
Method: Least Squares				
Date: 01/27/09 Time: 15:48				
Sample: 1978Q1 1985Q4				
Included observations: 32				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1160.000	59.99041	19.33642	0.0000
DUM1	62.12500	84.83926	0.732267	0.4701
DUM2	307.5000	84.83926	3.624501	0.0011
DUM3	409.7500	84.83926	4.829722	0.0000
R-squared	0.531797	Mean dependent var	1354.844	
Adjusted R-squared	0.481632	S.D. dependent var	235.6719	
S.E. of regression	169.6785	Akaike info criterion	13.22216	
Sum squared resid	806142.4	Schwarz criterion	13.40537	
Log likelihood	-207.5545	F-statistic	10.60102	
Durbin-Watson stat	0.392512	Prob(F-statistic)	0.000079	

Câu hỏi

- 1) Viết hàm SRF.
- 2) Tính số tủ lạnh bán được trung bình trong các quý.
- 3) So sánh số tủ lạnh bán được trong các quý. Giải thích.
- 4) Kiểm định giả thiết cho rằng số tủ lạnh bán được trong quý 1 và quý 4 là như nhau.

42

Bài Giảng
KINH TẾ LƯỢNG
 (Econometric)

**Chương 3. Kiểm Định
 Giả Thuyết Mô Hình**

GV: ThS. Nguyễn Trung Đông
 nguyentrungdong144@yahoo.com

**Chương 3. Kiểm Định
 Giả Thuyết Mô Hình**

Ba giả thiết quan trọng của mô hình hồi quy tuyến tính là

a) Các sai số ngẫu nhiên ε_i trong hàm hồi quy tổng thể có *phương sai không đổi* và bằng σ^2 .

b) Không có hiện tượng *cộng tuyến* giữa các biến giải thích.

c) Không có hiện tượng *tự tương quan* giữa các nhiễu.

**Chương 3. Kiểm Định
 Giả Thuyết Mô Hình**

- ❖ Phương sai thay đổi
- ❖ Đa cộng tuyến
- ❖ Tự tương quan

1. Phương sai thay đổi

➤ Xét mô hình hồi quy trong đó giả thiết a) bị vi phạm, nghĩa là khi phương sai của các nhiễu ε_i là σ_i^2 (thay đổi theo từng quan sát một).

➤ Khi đó phương pháp OLS dùng để ước lượng các hệ số hồi quy được thay đổi, cụ thể ta xét hai phương pháp.

- a. Phương pháp OLS có trọng số
- b. Phương pháp OLS tổng quát

1.1. Phương pháp OLS có trọng số

Xét hàm hồi quy tuyến tính: $Y = \beta_1 + \beta_2 X + \varepsilon$

Giá trị quan sát thứ i của Y có dạng

$$Y_i = \beta_1 + \beta_2 X_i + \varepsilon_i, \quad i = \overline{1, n}$$

Trong đó ε_i là sai số ngẫu nhiên ở quan sát thứ i và $\text{var}(\varepsilon_i) = \sigma_i^2$

$$\text{Đặt } w_i = \begin{cases} \frac{1}{\sigma_i^2}, & \text{nếu } \sigma_i^2 \text{ đã biết} \\ \frac{1}{X_i}, & \text{nếu } \sigma_i^2 \text{ chưa biết} \end{cases}, \quad i = \overline{1, n}$$

Ta gọi $w_i, i = \overline{1, n}$ là trọng số

5

1.1. Phương pháp OLS có trọng số

❖ Khi đó tìm hàm hồi quy mẫu có dạng

$$Y^* = \beta_1^* + \beta_2^* X$$

❖ Giả sử quan sát thứ i của Y^* có dạng

$$Y_i^* = \beta_1^* + \beta_2^* X_i, \quad i = \overline{1, n}$$

❖ Phần dư ở quan sát thứ i có dạng

$$e_i = Y_i - Y_i^* = Y_i - \beta_1^* - \beta_2^* X_i, \quad i = \overline{1, n}$$

❖ Tìm (β_1^*, β_2^*) sao cho

$$f(\beta_1^*, \beta_2^*) = \sum_{i=1}^n w_i e_i^2 \rightarrow \min$$

6

1.1. Phương pháp OLS có trọng số

❖ Hàm số đạt cực trị khi

$$\frac{\partial f(\beta_1^*, \beta_2^*)}{\partial \beta_i^*} = 0$$

❖ Từ đó ta có hệ phương trình

$$\begin{cases} \left(\sum_{i=1}^n w_i \right) \beta_1^* + \left(\sum_{i=1}^n w_i X_i \right) \beta_2^* = \sum_{i=1}^n w_i Y_i \\ \left(\sum_{i=1}^n w_i X_i \right) \beta_1^* + \left(\sum_{i=1}^n w_i X_i^2 \right) \beta_2^* = \sum_{i=1}^n w_i X_i Y_i \end{cases}$$

❖ Hệ PT trên luôn có nghiệm (β_1^*, β_2^*) .

7

1.2. Phương pháp OLS tổng quát

❖ Xét hàm hồi quy tuyến tính

$$Y = \beta_1 + \beta_2 X + \varepsilon$$

❖ Giá trị quan sát thứ i của Y có dạng

$$Y_i = \beta_1 + \beta_2 X_i + \varepsilon_i, \quad i = \overline{1, n}$$

❖ Trong đó ε_i là sai số ngẫu nhiên ở quan sát thứ i và $\text{var}(\varepsilon_i) = \sigma_i^2$

❖ Chia 2 vế cho σ_i ($\sigma_i > 0$), ta được

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

8

1.2. Phương pháp OLS tổng quát

❖ Đặt

$$Y_i^* = \frac{Y_i}{\sigma_i}, X_{0,i}^* = \frac{1}{\sigma_i}, X_i^* = \frac{X_i}{\sigma_i}, \varepsilon_i^* = \frac{\varepsilon_i}{\sigma_i}$$

❖ Đẳng thức trên được viết lại thành

$$Y^* = \beta_1 X_0^* + \beta_2 X^* + \varepsilon^*$$

❖ Chú ý khi đó

$$\text{var}(\varepsilon_i^*) = \frac{\text{var}(\varepsilon_i)}{\sigma_i^2} = 1$$

9

1.3. Nguyên nhân của phương sai thay đổi

- Do bản chất mối quan hệ trong kinh tế chứa đựng hiện tượng này.
- Do kỹ thuật thu nhập số liệu được cải tiến, sai lầm phạm phải ít đi.
- Do con người học được hành vi trong quá khứ.
- Do trong mẫu có giá trị bất thường.

10

1.4. Hậu quả của phương sai thay đổi

- Các ước lượng OLS vẫn là ước lượng tuyến tính, không chệch nhưng không phải là ước lượng hiệu quả.
- Ước lượng của phương sai bị chệch. Do đó, các kiểm định Student và Fisher không còn đáng tin cậy nữa.
- Kết quả dự báo không hiệu quả khi sử dụng các ước lượng OLS.

11

1.5. Phát hiện PSTĐ

1.5.1. Xét đồ thị phần dư

Ví dụ. Cho các số liệu về chi tiêu cho tiêu dùng (Y) và thu nhập (X) hàng tháng của 20 hộ gia đình ở một vùng nông thôn (đơn vị 10.000 đ)

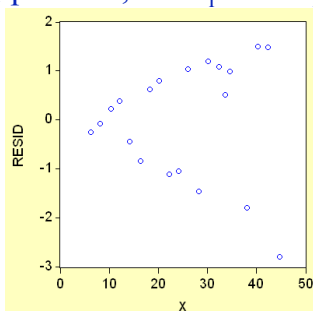
STT	X	Y	STT	X	Y
1	22.3	19.9	12	34.5	33.1
2	32.3	31.2	13	38	33.5
3	33.6	31.8	14	14.1	13.1
4	12.1	12.1	15	16.4	14.8
5	42.3	40.7	16	24.1	21.6
6	6.2	6.1	17	30.1	29.3
7	44.7	38.6	18	28.3	25
8	26.1	25.5	19	18.2	17.9
9	10.3	10.3	20	20.1	19.8
10	40.2	38.8	19	18.2	17.9
11	8.1	8	20	20.1	19.8

12

Khi đó, ta tìm được mô hình hồi quy sau

$$\hat{Y} = 0.7075 + 0.9103X; R^2 = 0.9878$$

và đồ thị phần dư, của e_i theo X_i



13

1.5.2. Kiểm định Park

➤ Park đã hình thức hóa phương pháp đồ thị cho rằng σ_i^2 là một hàm theo X dạng đề nghị là

$$\sigma_i^2 = \sigma^2 X_i^{\beta_2} e^{\varepsilon_i}$$

➤ Lấy logarit 2 vế ta được

$$\ln \sigma_i^2 = \ln \sigma^2 + \beta_2 \ln X_i + \varepsilon_i$$

➤ Trong đó ε_i là sai số ngẫu nhiên.

14

1.5.2. Kiểm định Park

➤ Do σ_i^2 chưa biết nên Park đề nghị dùng e_i^2 thay cho σ_i^2 và ước lượng hồi quy sau

$$\ln e_i^2 = \ln \sigma^2 + \beta_2 \ln X_i + \varepsilon_i = \beta_1 + \beta_2 \ln X_i + \varepsilon_i$$

➤ Trong đó $\beta_1 = \ln \sigma^2$ và e_i^2 tính từ hồi quy gốc.

➤ Các bước của kiểm định Park gồm:

15

1.5.2. Kiểm định Park

Bước 1: Ước lượng hàm hồi quy gốc cho dù có hiện tượng phương sai thay đổi.

Bước 2: Tính e_i^2 , $\ln e_i^2$, $\ln X_i$

Bước 3: Ước lượng hàm MH hồi quy

$$\ln e_i^2 = \beta_1 + \beta_2 \ln X_i + \varepsilon_i$$

Bước 4: KĐ giả thuyết $H_0: \beta_2 = 0$ tức là “không có hiện tượng phương sai thay đổi”

❖ Nếu bác bỏ H_0 , nghĩa là có hiện tượng phương sai thay đổi.

16

Ước lượng MH: $\ln e_i^2 = \beta_1 + \beta_2 \ln X_i + \varepsilon_i$

Dependent Variable: LOG(RESID*2)

Method: Least Squares

Date: 06/08/12 Time: 21:09

Sample: 1 20

Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-8.529469	1.038973	-8.209520	0.0000
LOG(X)	2.581552	0.330972	7.799916	0.0000

R-squared	0.771686	Mean dependent var	-0.553733
Adjusted R-squared	0.759002	S.D. dependent var	1.676648
S.E. of regression	0.823093	Akaike info criterion	2.543145
Sum squared resid	12.19468	Schwarz criterion	2.642718
Log likelihood	-23.43145	F-statistic	60.83870
Durbin-Watson stat	1.516201	Prob(F-statistic)	0.000000

1.5.3. Kiểm định Gleiser

Tương tự như kiểm định Park, sau khi thu được các phần dư e_i^2 , Gleiser đề nghị dùng

$$|e_i| = \beta_1 + \beta_2 X_i + \varepsilon_i$$

$$|e_i| = \beta_1 + \beta_2 \sqrt{X_i} + \varepsilon_i$$

$$|e_i| = \beta_1 + \beta_2 \frac{1}{X_i} + \varepsilon_i$$

$$|e_i| = \beta_1 + \beta_2 \frac{1}{\sqrt{X_i}} + \varepsilon_i$$

Trong đó ε_i là sai số ngẫu nhiên

18

Ước lượng MH 1: $|e_i| = \beta_1 + \beta_2 X_i + \varepsilon_i$

Dependent Variable: ABS(RESID)

Method: Least Squares

Date: 06/12/12 Time: 22:00

Sample: 1 20

Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.166276	0.191084	-0.870170	0.3957
X	0.045755	0.006923	6.608966	0.0000

R-squared	0.708164	Mean dependent var	0.982184
Adjusted R-squared	0.691951	S.D. dependent var	0.640368
S.E. of regression	0.355419	Akaike info criterion	0.863598
Sum squared resid	2.273802	Schwarz criterion	0.963171
Log likelihood	-6.635981	F-statistic	43.67843
Durbin-Watson stat	1.781207	Prob(F-statistic)	0.000003

Ước lượng MH 2: $|e_i| = \beta_1 + \beta_2 \sqrt{X_i} + \varepsilon_i$

Dependent Variable: ABS(RESID)

Method: Least Squares

Date: 06/12/12 Time: 22:03

Sample: 1 20

Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.070759	0.338661	-3.161740	0.0054
SQR(X)	0.422536	0.067597	6.250797	0.0000

R-squared	0.684611	Mean dependent var	0.982184
Adjusted R-squared	0.667090	S.D. dependent var	0.640368
S.E. of regression	0.369482	Akaike info criterion	0.941211
Sum squared resid	2.457307	Schwarz criterion	1.040784
Log likelihood	-7.412107	F-statistic	39.07246
Durbin-Watson stat	1.809645	Prob(F-statistic)	0.000007

Ước lượng MH 3: $|e_i| = \beta_1 + \beta_2 \frac{1}{X_i} + \varepsilon_i$

Dependent Variable: ABS(RESID)
 Method: Least Squares
 Date: 06/12/12 Time: 22:04
 Sample: 1 20
 Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.647269	0.187054	8.806400	0.0000
1/X	-12.30281	2.880094	-4.271668	0.0005
R-squared	0.503409	Mean dependent var	0.982184	
Adjusted R-squared	0.475821	S.D. dependent var	0.640368	
S.E. of regression	0.463628	Akaike info criterion	1.395172	
Sum squared resid	3.869120	Schwarz criterion	1.494745	
Log likelihood	-11.95172	F-statistic	18.24715	
Durbin-Watson stat	1.610693	Prob(F-statistic)	0.000459	

Ước lượng MH 4: $|e_i| = \beta_1 + \beta_2 \frac{1}{\sqrt{X_i}} + \varepsilon_i$

Dependent Variable: ABS(RESID)
 Method: Least Squares
 Date: 06/12/12 Time: 22:06
 Sample: 1 20
 Included observations: 20

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.532172	0.325954	7.768500	0.0000
1/SQR(X)	-6.972226	1.401908	-4.973384	0.0001
R-squared	0.578795	Mean dependent var	0.982184	
Adjusted R-squared	0.555395	S.D. dependent var	0.640368	
S.E. of regression	0.426989	Akaike info criterion	1.230525	
Sum squared resid	3.281761	Schwarz criterion	1.330098	
Log likelihood	-10.30525	F-statistic	24.73455	
Durbin-Watson stat	1.659035	Prob(F-statistic)	0.000098	

1.5.4. Kiểm định White

White không đòi ε phải có phân phối chuẩn. Đây là một kiểm định tổng quát về sự thuần nhất của phương sai.

Xét mô hình hồi quy ba biến sau

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

23

1.5.4. Kiểm định White

➤B1: Ước lượng và thu được các phần dư e_i^2

➤B2: Ước lượng mô hình

$$e_i^2 = \alpha_1 + \alpha_2 X_2 + \alpha_3 X_3 + \alpha_4 X_2^2 + \alpha_5 X_3^2 + \alpha_6 X_2 X_3 + V$$

Trong đó phải có hệ số chặn. Xét hệ số xác định R^2 của mô hình này.

➤B3: H_0 "Phương sai của sai số ngẫu nhiên không đổi"

➤B4: Nếu $nR^2 > \chi_\alpha^2(k-1)$, bác bỏ GT H_0 .

24

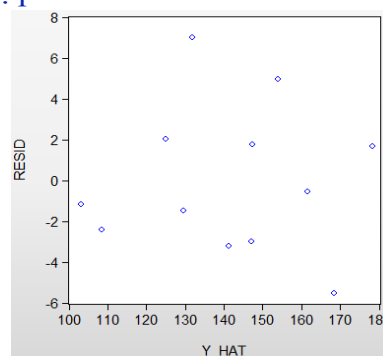
Kiểm định White có thể mở rộng cho mô hình hồi quy có k biến bất kỳ

Ví dụ: Bảng sau cho biết số liệu về doanh thu (Y), chi phí quảng cáo (X2), tiền lương của nhân viên tiếp thị (X3) của 12 công nhân (đơn vị là triệu đồng).

STT	1	2	3	4	5	6	7	8	9	10	11	12
Y_i	127	149	106	163	102	180	161	128	139	144	159	138
X_{2i}	18	25	19	24	15	26	25	16	17	23	22	15
X_{3i}	10	11	6	16	7	17	14	12	12	12	14	15

25

Hàm SRF: $\hat{Y} = 32.28 + 2.51X_2 + 4.76X_3$
Đồ thị phần dư



26

1.5.4. Kiểm định White

➤ Nhìn vào đồ thị phần dư ta thấy độ rộng của phần dư không thay đổi khi \hat{Y} tăng. Vậy khả năng mô hình không xảy ra hiện tượng phương sai thay đổi. Dùng kiểm định White, phát biểu GT

H_0 : Phương sai của sai số NN không đổi

p_value = 0.4215 > α

27

Heteroskedasticity Test: White

F-statistic	0.843820	Prob. F(5,6)	0.5645
Obs*R-squared	4.954369	Prob. Chi-Square(5)	0.4215
Scaled explained SS	2.033496	Prob. Chi-Square(5)	0.8445

Test Equation:
Dependent Variable: RESID^2
Method: LeastSquares
Date: 07/28/13 Time: 16:21
Sample: 1 12
Included observations: 12

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-302.2919	201.0729	-1.503394	0.1834
X2	31.71005	20.36090	1.557400	0.1704
X2^2	-0.879882	0.520465	-1.690568	0.1419
X2*X3	0.208625	0.536312	0.385270	0.7133
X3	0.587451	10.18193	0.057695	0.9559
X3^2	-0.075395	0.534322	-0.141104	0.8924

R-squared	0.412664	Mean dependent var	12.01991
Adjusted R-squared	-0.076416	S.D. dependent var	15.16494
S.E. of regression	15.73369	Akaike info criterion	8.656339
Sum squared resid	1495.295	Schwarz criterion	8.896792
Log likelihood	-45.93803	Hannan-Quinn criter	8.566574
F-statistic	0.843820	Durbin-Watson stat	2.579780
Prob(F-statistic)	0.564521		

28

1.6. Biện pháp khắc phục

Có 2 cách xử lý :

- Khi biết σ_i^2 , ta có thể dùng phương pháp bình phương nhỏ nhất có trọng số đã trình bày ở trên.
- Khi chưa biết σ_i^2 , ta cần thêm những giả thuyết nhất định về σ_i^2 và biến đổi mô hình hồi quy gốc về mô hình mà phương sai không đổi.

29

2. Đa cộng tuyến

2.1. Định Nghĩa

2.2. Hậu quả

2.3. Phát hiện đa cộng tuyến

2.4. Khắc phục

30

Đa cộng tuyến là gì ?

- Theo giả thiết của phương pháp OLS thì các biến độc lập trong mô hình không có mối tương quan tuyến tính với nhau.
- Nếu giả thiết này vi phạm thì mô hình sẽ xảy ra hiện tượng đa cộng tuyến.
- Như vậy, đa cộng tuyến là hiện tượng mà các biến độc lập trong mô hình phụ thuộc tuyến tính với nhau dưới dạng hàm số.

31

2.1. Định nghĩa

- Xét hàm hồi quy tuyến tính k biến

$$Y = \beta_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon$$

- Nếu tồn tại $\lambda_2, \lambda_3, \dots, \lambda_k$ không đồng thời bằng 0 sao cho

$$\lambda_2 X_2 + \lambda_3 X_3 + \dots + \lambda_k X_k = 0$$

- Ta nói giữa các biến X_i ($i = 2, 3, \dots, k$) xảy ra hiện tượng đa cộng tuyến hoàn hảo.

2.1. Định nghĩa

➤ Nếu $\lambda_2 X_2 + \lambda_3 X_3 + \dots + \lambda_k X_k + V = 0$

Với V là một sai số ngẫu nhiên, ta nói giữa các biến X_i ($i = 2, 3, \dots, k$) xảy ra hiện tượng đa cộng tuyến không hoàn hảo.

33

Ví dụ. Giả sử ta có bảng dữ liệu giả định như sau

X_2	10	15	18	24	30
X_3	50	75	90	120	150
X_3^*	52	75	97	129	152
e	2	0	7	9	2

Ta thấy rằng

$$X_3^* = X_3 + e; X_3 = 5X_2; X_3^* = 5X_2 + e$$

$$r_{X_2 X_3} = 1; r_{X_2 X_3^*} \approx 0.996$$

34

2.2. Hậu quả

Các hệ số hồi quy không xác định

Thì

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n Y_i X_{2,i} \cdot \lambda^2 \sum_{i=1}^n X_{2,i}^2 - \lambda \sum_{i=1}^n Y_i X_{2,i} \cdot \lambda \sum_{i=1}^n X_{2,i}^2}{\sum_{i=1}^n X_{2,i}^2 \cdot \lambda^2 \sum_{i=1}^n X_{2,i}^2 - \lambda^2 \left(\sum_{i=1}^n X_{2,i}^2 \right)^2} = \frac{0}{0}$$

Do đó, nếu $x_{3,i} = \lambda x_{2,i}$,

với λ là một hằng số khác 0

35

Phương sai và hiệp phương sai lớn

$$\text{var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum_{i=1}^n X_{2,i}^2 (1 - r_{2,3}^2)}$$

$$\text{var}(\hat{\beta}_3) = \frac{\sigma^2}{\sum_{i=1}^n X_{3,i}^2 (1 - r_{2,3}^2)}$$

$$\text{cov}(\hat{\beta}_2, \hat{\beta}_3) = \frac{-r_{2,3} \sigma^2}{(1 - r_{2,3}^2) \sqrt{\sum_{i=1}^n X_{2,i}^2 \sum_{i=1}^n X_{3,i}^2}}$$

V_i - Nên khoảng tin cậy rộng hơn $\sqrt{\text{var}(\hat{\beta}_3)}$

Tỷ số $t = \frac{\hat{\beta}_2}{\text{se}(\hat{\beta}_2)}$ không có ý nghĩa

- Tỷ số $t = \frac{\hat{\beta}_2}{\text{se}(\hat{\beta}_2)}$ không có ý nghĩa

Do đó nếu ta dùng tỷ số $t = \frac{\hat{\beta}_2}{\text{se}(\hat{\beta}_2)}$

để kiểm định giả thiết

$$H: \beta_2 = 0$$

nên khi có hiện tượng cộng tuyến cao thì sai số chuẩn sẽ rất lớn làm cho giá trị t nhỏ đi, kết quả làm tăng khả năng chấp nhận H .

37

2.2. Hậu quả

➤ R^2 Cao nhưng tỷ số t nhỏ.

➤ Dấu của các hệ số hồi có thể sai.

➤ Thêm vào hay bớt đi các biến cộng tuyến với các biến khác, mô hình sẽ thay đổi về dấu hoặc thay đổi về độ lớn của các ước lượng.

❖ **Tóm lại:** Dấu hiệu chủ yếu của đa cộng tuyến là làm tăng sai số chuẩn.

2.3. Phát hiện đa cộng tuyến

1) Hệ số R^2 lớn nhưng tỷ số t nhỏ

➤ Dấu hiệu này thể hiện nghịch lý là các hệ số hồi quy có thể sai so với thực tế.

➤ **Nhược điểm:** Chỉ thể hiện rõ khi có đa cộng tuyến ở mức độ cao.

39

2.3. Phát hiện đa cộng tuyến

2) Hệ số tương quan $r_{X,Z}$ giữa các cặp biến giải thích cao: Hệ số này được tính bởi

$$r_{X,Z} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Z_i - \bar{Z})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Z_i - \bar{Z})^2}}$$

Lưu ý: Ta có thể dùng ma trận tương quan Theo Kennedy, nếu hệ số tương quan từ 0.8 trở lên thì đa cộng tuyến trở lên nghiêm trọng.

2.3. Phát hiện đa cộng tuyến

- 3) Dùng mô hình hồi quy phụ
- Hồi quy của mỗi biến độc lập theo các biến độc lập còn lại. Tính R_j^2 và F
 - Kiểm định giả thuyết $H_0: R_j^2 = 0$
 - Nếu ta bác bỏ H_0 thì mô hình xảy ra hiện tượng đa cộng tuyến.

41

2.3. Phát hiện đa cộng tuyến

- 4) Dùng nhân tử phóng đại phương sai (VIF)

$$VIF_j = \frac{1}{1 - R_j^2}$$

Trong đó R_j^2 là hệ số xác định của mô hình hồi quy phụ.

Nếu $VIF_j > 10$ thì X_j có đa cộng tuyến cao với các biến giải thích khác.

42

2.4. Khắc phục

- Loại trừ một biến độc lập ra khỏi mô hình.
- **Bước 1.** Khảo sát cặp biến giải thích nào có liên quan chặt chẽ.
- **Bước 2.** Tính R^2 đối với các hàm hồi quy: có mặt cả hai biến độc lập, không có mặt một trong hai biến độc lập.
- **Bước 3.** Loại biến nào mà giá trị R^2 tính được khi không có mặt biến độc lập đó là lớn hơn.

43

2.4. Khắc phục

- Thu thập thêm số liệu hoặc lấy mẫu mới

$$\text{Var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum_{i=1}^n x_{2,i}^2 (1 - r_{2,3}^2)}$$

Khi cỡ mẫu tăng, $\sum_{i=1}^n x_{2,i}^2$ nói chung sẽ tăng và khi đó với $r_{2,3}$ cho trước, phương sai của $\hat{\beta}_2$ sẽ giảm và điều này giúp ta ước lượng β_2 chính xác hơn.

44

Ví dụ. Khảo sát chi tiêu cho tiêu dùng, thu nhập và sự giàu có ta có bảng số liệu sau

Y	70	65	90	95	110	115	120	140	155	150
X ₂	80	100	120	140	160	180	200	220	240	260
X ₃	810	1009	1273	1425	1633	1876	2052	2201	2435	2686

Trong đó

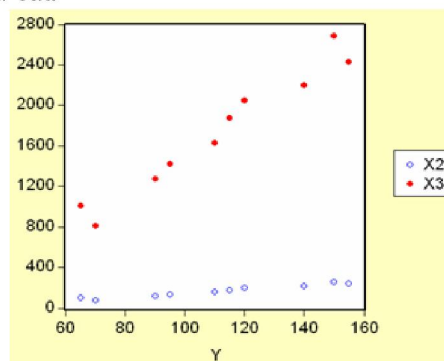
Y : Chi tiêu cho tiêu dùng (\$)

X₂ : Thu nhập (\$)

X₃ : Sự giàu có (\$)

45

Ta có đồ thị phân tán của Y theo từng biến X₂ và X₃ như sau



46

Ta có kết quả hồi quy như sau

Variable	Coefficient	Std. Error	t - Statistic	Prob.
C	24.77473	6.752500	3.668972	0.0080
X ₂	0.941537	0.822898	1.144172	0.2902
X ₃	-0.042435	0.080664	-0.526062	0.6151
R - squared	0.963504	Adjusted R -squared	0.953077	

Ma trận tương quan giữa các biến

Correlation Matrix			
	X2	X3	Y
X2	1.000000	0.998962	0.980847
X3	0.998962	1.000000	0.978100
Y	0.980847	0.978100	1.000000

47

- Xét các mô hình hồi quy của Y theo từng biến :

+ Hồi quy Y theo X₂, ta có kết quả

$$\hat{Y} = 24.45455 + 0.509091X_2 ; R^2 = 0.962062$$

+ Hồi quy Y theo X₃,

$$\hat{Y} = 24.41104 + 0.049764X_3 ; R^2 = 0.956679$$

- Xét mô hình hồi quy phụ của X₃ theo X₂

$$X_3 = 7.545455 + 10.19091X_2 ; R^2 = 0.997926$$

- Ngoài ra ta có giá trị

$$VIF = \frac{1}{1 - r_{23}^2} = \frac{1}{1 - 0.997926} = 482.16901$$

48

Kết luận : X_2, X_3 Có tương quan tuyến tính khá cao ($r_{2,3} = 0,998962$). Do đó mô hình trên xảy ra hiện tượng đa cộng tuyến. Vậy để hạn chế hậu quả của đa cộng tuyến trong mô hình trên, ta có thể giải quyết theo phương pháp loại trừ biến độc lập.

49

3. Tự tương quan

3.1. Nguyên nhân

3.2. Hậu quả

3.3. Phát hiện tự tương quan

50

3.1. Nguyên nhân

i) Một số nguyên nhân khách quan

Quán tính: Các chuỗi thời gian như: tổng sản lượng, chỉ số giá, thất nghiệp... mang tính chu kỳ. Khi đó các quan sát kế tiếp có nhiều khả năng phụ thuộc vào nhau.

Hiện tượng mạng nhện: là hiện tượng một biến cần một thời gian trễ để phản ứng lại với sự thay đổi của biến khác.

3.1. Nguyên nhân

i) Một số nguyên nhân khách quan

Các độ trễ: Trong chuỗi thời gian, ta gặp hiện tượng biến phụ thuộc ở thời kỳ t phụ thuộc vào chính nó ở thời kỳ $t-1$ và các biến khác.

Ví dụ: Mô hình sau

$$Y_t = \beta_1 + \beta_2 X_t + Y_{t-1} + \varepsilon_t$$

52

3.1. Nguyên nhân

ii) Một số nguyên nhân chủ quan

- ❖ Xử lý số liệu: do việc “làm trơn” số liệu → loại bỏ những quan sát “gai góc”.
- ❖ Sai lệch do mô hình: Bỏ sót biến, dạng hàm sai.

53

3.2. Hậu quả

- Các ước lượng OLS vẫn là ước lượng tuyến tính, không chệch nhưng không còn hiệu quả nữa.
- Ước lượng của phương sai bị chệch nên kiểm định t và F không hiệu quả nữa.
- Thường R^2 được ước lượng khá cao so với giá trị thực.
- Sai số chuẩn của các giá trị dự báo không còn tin cậy nữa.

54

3.3. Phát hiện tự tương quan

3.3.1. Phương pháp đồ thị

- Ta vẽ đồ thị của phần dư e_i theo e_{i-1} . Nếu e_i đồng biến theo e_{i-1} . Thì ta kết luận có hiện tượng tự tương quan

55

3.3. Phát hiện tự tương quan

3.3.2. Kiểm định d của Durbin - Watson

Thống kê d của Durbin – Watson xác định bởi

$$d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2} \approx 2(1 - \hat{\rho})$$

56

3.3.2. Kiểm định d của Durbin - Watson

Trong đó

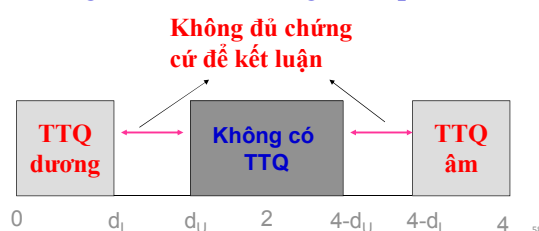
$$\hat{\rho} = \frac{\sum_{t=2}^n e_t e_{t-1}}{\sum_{t=1}^n e_t^2}$$

là một ước lượng của hệ số tương quan ρ . Khi đó ta có thể dùng bảng kết quả để kiểm tra vấn đề tự tương quan.

57

3.3.2. Kiểm định d của Durbin - Watson

Kiểm định Durbin Watson, dùng trong trường hợp tự tương quan bậc nhất; Không có giá trị trễ của biến phụ thuộc là biến giải thích; Không mất quan sát.



58

3.3.2. Kiểm định d của Durbin - Watson

Trong thực tế, kiểm định Durbin Watson, người ta sử dụng quy tắc sau

- Nếu $0 < d < 1$ Mô hình có tự tương quan dương.
- Nếu $1 < d < 3$ Mô hình không có tự tương quan.
- Nếu $3 < d < 4$ Mô hình có tự tương quan âm.

59

Ví dụ. Cho các số liệu về nhu cầu về kem (Y đơn vị : pints = 0.473 lít) và thu nhập hàng tuần của gia đình (X đơn vị USD).

Y	X	Y	X	Y	X	Y	X
0.386	78	0.381	82	0.269	76	0.326	92
0.374	79	0.47	80	0.256	79	0.309	95
0.393	81	0.443	78	0.286	82	0.359	96
0.425	80	0.386	84	0.298	85	0.376	94
0.406	76	0.342	86	0.329	86	0.416	96
0.344	78	0.319	85	0.318	83	0.437	91
0.327	82	0.307	87	0.381	84	0.548	90
0.288	79	0.284	94				

60

Hồi quy Y theo X ta được kết quả

Dependent Variable: Y
Method: Least Squares
Date: 02/29/08 Time: 14:34
Sample: 1 30
Included observations: 30

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.316715	0.168665	1.877771	0.0709
X	0.000505	0.001988	0.253942	0.8014
R-squared	0.002298	Mean dependent var	0.359433	
Adjusted R-squared	-0.033334	S.D. dependent var	0.065791	
S.E. of regression	0.066878	Akaike info criterion	-2.507551	
Sum squared resid	0.125235	Schwarz criterion	-2.414138	
Log likelihood	39.61326	F-statistic	0.064487	
Durbin-Watson stat	0.392752	Prob(F-statistic)	0.801396	

3.3. Phát hiện tự tương quan

3.3.3. Kiểm định Breusch - Godfrey (BG)

➤ Xét mô hình hồi quy: $Y_i = \beta_1 + \beta_2 X_i + e_i$

Trong đó các e_i có tự tương quan bậc p

$$e_i = \rho_1 e_{i-1} + \rho_2 e_{i-2} + \dots + \rho_p e_{i-p} + \varepsilon_i$$

ε_i thỏa mãn các giả thuyết của OLS

➤ Với giả thuyết $H_0 : \rho_1 = \rho_2 = \dots = \rho_p = 0$

(không có tự tương quan bậc p)

62

3.3.3. Kiểm định Breusch - Godfrey (BG)

Các bước tiến hành kiểm định BG như sau

➤ B1: Ước lượng mô hình: $Y_i = \beta_1 + \beta_2 X_i + e_i$
bằng phương pháp OLS.

➤ B2: Ước lượng mô hình:

$$e_i = \beta_1 + \beta_2 X_i + \rho_1 e_{i-1} + \rho_2 e_{i-2} + \dots + \rho_p e_{i-p} + V_i$$

Từ kết quả của ước lượng ta tính được R^2

➤ B3: Với n đủ lớn, ta có $(n-p)R^2 \sim \chi^2(p)$

➤ B4: Nếu $(n-p)R^2 > \chi^2_\alpha(p)$ bác bỏ H_0 .

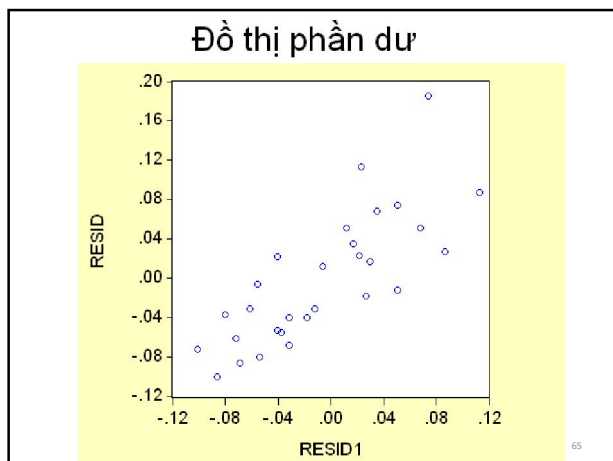
63

Với ví dụ trên: Hồi quy của Y theo X

Dependent Variable: Y
Method: Least Squares
Date: 02/29/08 Time: 14:34
Sample: 1 30
Included observations: 30

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.316715	0.168665	1.877771	0.0709
X	0.000505	0.001988	0.253942	0.8014
R-squared	0.002298	Mean dependent var	0.359433	
Adjusted R-squared	-0.033334	S.D. dependent var	0.065791	
S.E. of regression	0.066878	Akaike info criterion	-2.507551	
Sum squared resid	0.125235	Schwarz criterion	-2.414138	
Log likelihood	39.61326	F-statistic	0.064487	
Durbin-Watson stat	0.392752	Prob(F-statistic)	0.801396	

64



Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	25.17764	Probability	0.000001	
Obs*R-squared	19.78460	Probability	0.000051	
Test Equation:				
Dependent Variable: RESID				
Method: Least Squares				
Date: 02/29/08 Time: 14:39				
Presample missing value lagged residuals set to zero.				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.132640	0.105564	-1.256495	0.2201
X	0.001626	0.001249	1.302498	0.2042
RESID(-1)	1.160705	0.217751	5.330431	0.0000
RESID(-2)	-0.259322	0.229175	-1.131547	0.2662
R-squared	0.659487	Mean dependent var	3.79E-17	
Adjusted R-squared	0.620196	S.D. dependent var	0.065715	
S.E. of regression	0.040499	Akaike info criterion	-3.451518	
Sum squared resid	0.042644	Schwarz criterion	-3.264692	
Log likelihood	55.77277	F-statistic	16.78509	
Durbin-Watson stat	2.007800	Prob(F-statistic)	0.000003	

Với ví dụ trên

- Tiến hành kiểm định tự tương quan bậc 2 bằng kiểm định BG, ta được kết quả sau:

$$p_value = 0.000051 < \alpha$$

- Bác bỏ H_0 , nghĩa là có tự tương quan bậc 2

- Nếu dùng KĐ Durbin Watson, ta có

$$\alpha = 0.05, n = 30, k' = 1, d_L = 1.35; d_U = 1.49$$

- Kết quả HQ của Y theo X, ta có giá trị

$d = 0.392752$ tức là $0 < d < d_L$, nên mô hình có tự tương quan bậc 1.

67

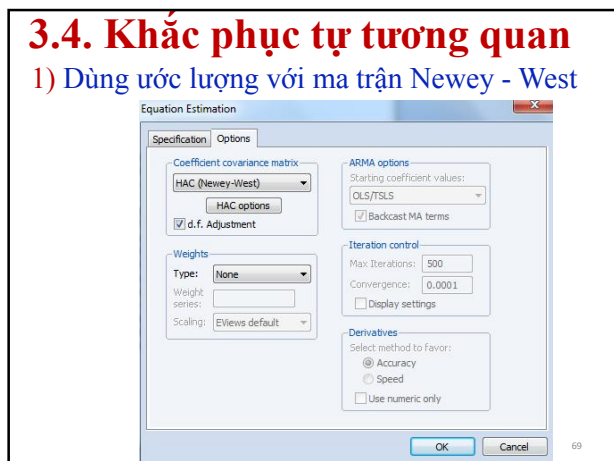
3.4. Khắc phục tự tương quan

- 1) Dùng ước lượng với ma trận Newey - West
- 2) Dùng GLS (Generalized Least Squares)
- 3) Các mô hình chuyên dùng cho dãy số thời gian (Kinh tế lượng nâng cao)

68

3.4. Khắc phục tự tương quan

1) Dùng ước lượng với ma trận Newey - West



MỘT SỐ ĐỀ THAM KHẢO

ĐỀ 1

Câu 1. Giả sử có số liệu thống kê về lãi suất ngân hàng và tổng vốn đầu tư trên địa bàn tỉnh A qua 5 năm liên tiếp như sau :

Lãi suất ngân hàng (%)	10	12	15	18	20
Tổng vốn đầu tư (tỉ đồng)	50	48	40	37	35

a. Hãy ước lượng hàm hồi quy tuyến tính của tổng đầu tư theo lãi suất ngân hàng và nêu ý nghĩa của các hệ số hồi quy tìm được.

b. Tính hệ số xác định mô hình và giải thích ý nghĩa.

c. Tìm khoảng tin cậy của hệ số hồi quy tổng thể β_2 , với độ tin cậy 95%.

d. Dự báo giá trị trung bình của tổng vốn đầu tư khi lãi suất ngân hàng là 13%, với độ tin cậy 95%. Giải thích kết quả.

e. Với mức ý nghĩa 5%, hãy cho biết lãi suất ngân hàng có ảnh hưởng đến tổng vốn đầu tư không?

f. Hãy viết hàm hồi quy khi đơn vị tính của tổng vốn đầu tư là triệu đồng.

g. Ta có kết quả khi dùng kiểm định White như sau :

White Heteroskedasticity Test:			
F-statistic	1.030339	Probability	0.492529
Obs*R-squared	2.537357	Probability	0.281203

Có phương sai thay đổi trong mô hình không, tại sao?

Câu 2. Người ta cho rằng tổng vốn đầu tư (Y : tỉ đồng) không chỉ phụ thuộc vào lãi suất ngân hàng (X_2 : %) mà còn phụ thuộc vào tốc độ tăng trưởng GDP (X_3 : %). Với số liệu gồm có 20 quan sát, người ta ước lượng được mô hình sau :

$$\begin{aligned} \hat{Y} &= 40,815 - 1,012X_2 + 2,123X_3 & R^2 &= 0,901 \\ t &= (2,748) \quad (-2,842) \quad (3,485) \end{aligned}$$

a. Nêu ý nghĩa của các hệ số hồi quy riêng?

b. Tìm khoảng tin cậy của các hệ số hồi quy tổng thể với độ tin cậy 95%.

c. Với mức ý nghĩa 5%, mô hình trên có phù hợp không?

d. Tính hệ số xác định mô hình có hiệu chỉnh.

Đề 2

Câu 1. Bảng sau đây cho chuỗi thời gian về mức tiêu dùng (Y : đơn vị 100000 VNĐ) và thu nhập (X : đơn vị 100000 VNĐ). Tính theo đầu người và tính theo giá cố định năm 1980 trong thời kỳ 1971 – 1990 ở một khu vực :

Năm	Y	X	Năm	Y	X
1971	48,34	52,02	1981	52,17	63,36
1972	48,54	52,41	1982	60,84	67,42
1973	47,44	51,55	1983	60,73	67,86
1974	54,58	58,88	1984	76,04	83,39
1975	55,00	59,66	1985	76,42	84,26
1976	63,49	68,42	1986	69,34	77,41
1977	59,22	64,27	1987	61,75	70,08
1978	57,77	63,01	1988	68,78	77,44
1979	60,22	65,61	1989	67,07	75,79
1980	55,40	61,05	1990	72,94	81,89

Giả sử Y và X có quan hệ tuyến tính

- Hãy ước lượng hàm hồi quy của mức tiêu dùng phụ thuộc vào thu nhập và nêu ý nghĩa kinh tế của các hệ số hồi quy tìm được.
- Tính hệ số xác định mô hình và giải thích ý nghĩa.
- Tìm khoảng tin cậy của các hệ số hồi quy tổng thể, với độ tin cậy 95%.
- Dự báo giá trị trung bình và cá biệt của mức tiêu dùng khi thu nhập là 8 triệu đồng, với độ tin cậy 95%. Giải thích kết quả.
- Với mức ý nghĩa 5%, hãy cho biết khi thu nhập thay đổi có ảnh hưởng đến mức tiêu dùng không?

Câu 2. Người ta cho rằng chi tiêu cho mặt hàng A (Y : ngàn đồng/tháng) không chỉ phụ thuộc vào thu nhập của người tiêu dùng (X : triệu đồng/tháng) mà còn phụ thuộc vào giới tính của người đó (D = 1 nếu là nam; D = 0 nếu là nữ). Với số liệu của một mẫu có kích thước n = 20 người ta đã ước lượng mô hình sau :

$$\begin{aligned}\hat{Y} &= 0,07 + 0,332D + 0,164X - 0,098XD \\ t &= (1,947) \quad (6,608) \quad (11,658) \quad (-5,303) \\ R^2 &= 0,974\end{aligned}$$

- Hãy nêu ý nghĩa của các hệ số hồi quy.

- b. Hãy ước lượng các hệ số hồi quy, với độ tin cậy 95%.
- c. Nêu ý nghĩa hệ số xác định mô hình và kiểm định sự phù hợp của mô hình với mức ý nghĩa 5%.
- d. Chi tiêu về loại hàng A của nam và nữ có giống nhau hay không? Kết luận với mức ý nghĩa 1%.

Đề 3

Câu 1. Bảng sau cho biết số liệu về tổng thu nhập (X : tỷ USD) và mức thuế (Y : tỷ USD) của một Doanh nghiệp.

X	Y	X	Y
14,95	1,84	291,69	43,14
17,83	2,53	148,63	22,33
7,42	0,95	168,78	23,31
99,26	14,55	148,23	19,74
14,14	1,88	75,26	10,07
67,09	10,85	181,32	26,72

Biết rằng Y và X có quan hệ tuyến tính với nhau

- a. Hãy ước lượng hàm hồi quy của Y theo X. Giải thích ý nghĩa kinh tế của các hệ số hồi quy nhận được.
- b. Tính hệ số xác định mô hình và giải thích ý nghĩa của kết quả nhận được.
- c. Tính hệ số co giãn của Y theo X tại điểm (\bar{X}, \bar{Y}) và giải thích ý nghĩa kết quả nhận được.
- d. Tìm khoảng tin cậy cho các hệ số hồi quy tổng thể với độ tin cậy 95%.
- e. Tìm khoảng tin cậy cho phương sai nhiễu với mức ý nghĩa 5%.
- f. Với mức ý nghĩa 5%, hãy cho biết khi thu nhập thay đổi có ảnh hưởng đến mức thuế không?
- g. Với mức tổng thu nhập $X_0 = 170$, hãy dự báo giá trị trung bình và giá trị cá biệt của mức thuế với độ tin cậy 95%. Giải thích kết quả.

Câu 2. Khảo sát sự liên hệ giữa sản lượng (Y : đơn vị tấn/ha) theo phân bón hóa học (X_2 : đơn vị tấn/ha) và thuốc trừ sâu (X_3 : đơn vị lít/ha) bằng cách dựa vào kết quả của mô hình hồi qui bội được cho trong bảng sau.

Dependent Variable: SANLUONG				
Method: Least Squares				
Date: 06/24/08 Time: 21:27				
Sample: 1 10				
Included observations: 10				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	32.30041	1.444532	22.36047	0.0000
PHANHOAHOC	0.575788	0.221453	2.600049	0.0354
THUOCTRUSAU	1.202950	0.236743	5.081247	0.0014
R-squared	0.993534	Mean dependent var	57.10000	
Adjusted R-squared	0.991687	S.D. dependent var	13.56835	
S.E. of regression	1.237094	Akaike info criterion	3.506733	
Sum squared resid	10.71282	Schwarz criterion	3.597508	
Log likelihood	-14.53366	F-statistic	537.8282	
Durbin-Watson stat	2.149723	Prob(F-statistic)	0.000000	

- Viết hàm SRF. Giải thích ý nghĩa các hệ số hồi quy.
- Phân hóa học có ảnh hưởng đến sản lượng hay không? Kết luận với mức ý nghĩa 5%.
- Mô hình trên có phù hợp hay không? Kết luận với mức ý nghĩa 5%.
- Mô hình trên có xảy ra hiện tượng tự tương quan hay không?
- Dự báo (điểm) giá trị trung bình của sản lượng khi phân hóa học là 20 tấn/ha và thuốc trừ sâu 16 lít/ha.

ĐỀ 4

Câu 1. Khảo sát doanh số bán của một loại hàng Y (triệu đồng/tháng) và chi phí quảng cáo X (triệu đồng/tháng) ở một khu vực, người ta thu được bảng số liệu sau :

Y	52	53	55	56	56	58	60	64	68	70
X	3	3	4	5	5	6	7	7	6	8

Giả sử Y và X có quan hệ tuyến tính

- Hãy ước lượng hàm hồi quy của Y theo X và nêu ý nghĩa của các hệ số hồi quy tìm được.
- Tính hệ số xác định mô hình và giải thích ý nghĩa.
Tính hệ số co dẫn của Y theo X tại (\bar{X}, \bar{Y}) và nêu ý nghĩa.
- Tìm khoảng tin cậy cho các hệ số hồi quy tổng thể, với độ tin cậy 95%. Nêu ý nghĩa kết quả.
- Tìm khoảng tin cậy cho phương sai nhiễu với mức ý nghĩa 5%.

e. Dự báo giá trị trung bình của doanh số bán khi chi phí quảng cáo là 9 triệu đồng/tháng, với độ tin cậy 95%. Giải thích kết quả.

f. Với mức ý nghĩa 5%, hãy cho biết chi phí quảng cáo ảnh hưởng đến doanh số bán hay không?

g. Ta có kết quả khi dùng kiểm định White như sau :

White Heteroskedasticity Test:			
F-statistic	0.882593	Probability	0.455176
Obs*R-squared	2.013861	Probability	0.365339

Có phương sai thay đổi trong mô hình không, tại sao?

h. Ta có kết quả sau khi dùng kiểm định BG như sau :

Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	1.996871	Probability	0.200512
Obs*R-squared	2.219518	Probability	0.136276

Có tự tương quan bậc nhất trong mô hình không, tại sao? Kết luận với mức ý nghĩa 5%.

Câu 2. Với số liệu của mẫu trên nhưng thêm vào biến Z (Z = 0 : nếu khu vực bán ở nông thôn; Z = 1 : nếu khu vực bán ở thành phố). Ta có mô hình sau :

$$\hat{Y} = 42,88 + 3,1769X - 1,67Z$$

$$se = (3,875) \quad (0,6825) \quad (2,218)$$

$$R^2 = 0,756$$

- Nêu ý nghĩa của các hệ số hồi quy riêng?
- Tìm khoảng tin cậy cho các hệ số hồi quy tổng thể, với độ tin cậy 95%.
- Theo bạn thì khu vực bán có ảnh hưởng đến doanh số bán không? Kết luận với mức ý nghĩa 5%.
- Với mức ý nghĩa 5%, mô hình trên có phù hợp không?
- Bạn chọn mô hình ở câu 1 hay câu 2, tại sao?

ĐỀ 5

Câu 1. Bảng sau cho biết số liệu về thu nhập khả dụng (X : 100.000 VNĐ) và chi tiêu cho tiêu dùng (Y : 100.000 VNĐ).

X	Y	X	Y
65,61	60,22	84,26	76,42
61,05	55,41	77,41	69,34
63,36	57,17	70,08	61,75
67,42	60,84	77,44	68,78
67,86	60,73	75,79	67,07
83,39	76,04	81,89	72,94

Biết rằng Y và X có quan hệ tuyến tính với nhau.

- Hãy ước lượng hàm hồi quy của Y theo X. Giải thích ý nghĩa kinh tế của các hệ số hồi quy nhận được.
- Tính hệ số xác định mô hình và giải thích ý nghĩa của kết quả nhận được.
- Tính hệ số co giãn của Y theo X tại điểm (\bar{X}, \bar{Y}) và giải thích ý nghĩa kết quả nhận được.
- Tìm khoảng tin cậy cho các hệ số hồi quy tổng thể với độ tin cậy 95%.
- Tìm khoảng tin cậy cho phương sai nhiễu với độ tin cậy 95%.
- Với mức ý nghĩa 5%, hãy cho biết khi thu nhập thay đổi có ảnh hưởng đến chi tiêu không?
- Hãy dự báo giá trị trung bình và giá trị cá biệt của chi tiêu khi mức thu nhập khả dụng là 8,5 triệu đồng, với độ tin cậy 95%.

Câu 2. Người ta cho rằng chi tiêu cho mặt hàng A (Y : ngàn đồng/tháng) không chỉ phụ thuộc vào thu nhập của người tiêu dùng (X : triệu đồng/tháng) mà còn phụ thuộc vào giới tính của người đó (D = 1 nếu là nam; D = 0 nếu là nữ). Với số liệu của một mẫu có kích thước $n = 20$ người ta đã ước lượng mô hình sau :

$$\begin{aligned} \hat{Y} &= 96,458 + 38,92X - 8,415D - 6,525XD \\ se &= (33,228) \quad (11,312) \quad (4,207) \quad (1,812) \end{aligned}$$

- Hãy nêu ý nghĩa của các hệ số hồi quy.
- Hãy ước lượng các hệ số hồi quy, với độ tin cậy 95%.
- Chi tiêu về loại hàng A của nam và nữ có giống nhau hay không? Kết luận với mức ý nghĩa 1%.

ĐỀ 6

Câu 1. Cho các số liệu của Y (chi tiêu tiêu dùng cá nhân) và X (tổng sản phẩm quốc nội) trong các năm 1980-1991 của Hoa kỳ cho trong bảng sau: (đơn vị tính là tỉ đô la năm 1987)

Năm	Y	X	Năm	Y	X
1980	2447,1	3776,3	1986	2969,1	4404,5
1981	2476,9	3843,1	1987	3052,2	4539,9
1982	2503,7	3760,3	1988	3162,4	4718,6
1983	2619,4	3906,6	1989	3223,3	4838
1984	2746,1	4148,5	1990	3260,4	4877,5
1985	2865,8	4279,8	1991	3240,8	4821

Giả sử Y và X có quan hệ tuyến tính

- Hãy ước lượng hàm hồi quy của chi tiêu tiêu dùng cá nhân phụ thuộc vào tổng sản phẩm quốc nội và nêu ý nghĩa kinh tế của các hệ số hồi quy tìm được.
- Tính hệ số xác định mô hình và giải thích ý nghĩa.
- Tìm khoảng tin cậy của các hệ số hồi quy tổng thể, với độ tin cậy 95%.
- Dự báo giá trị trung bình và cá biệt của chi tiêu tiêu dùng cá nhân khi tổng sản phẩm quốc nội là 5000 tỉ đô la, với độ tin cậy 95%. Giải thích kết quả.
- Với mức ý nghĩa 5%, hãy cho biết khi tổng sản phẩm quốc nội thay đổi có ảnh hưởng đến chi tiêu tiêu dùng cá nhân không?

Câu 2. Khảo sát về mối liên hệ giữa mức tiêu dùng của hộ gia đình (Y) theo thu nhập (X_2) và tài sản có khả năng chuyển đổi cao (X_3) được cho trong các bảng kết quả sau

Dependent Variable Y		Included observations : 25		
Variable	Coefficient	Std. Error	t_Statistic	Prob.
C	33,87971	19,11513	1,772403	0,0902
X_2	-26,00263	34,95898	-0,743804	0,4649
X_3	6,709261	8,74055	0,767602	0,4509
R-squared		0,741695	F-statistic	31,58532
Adjusted R-squared		0,718213	Prob(F-statistic)	0,000

Correlation Matrix				Coefficient Covariance Matrix			
	X ₂	X ₃	Y		C	X ₂	X ₃
X ₂	1,00000	0,999995	0,857191	C	365,3883	128,3093	-32,52119
X ₃	0,999995	1,00000	0,857438	X ₂	128,3093	1222,13	-305,5593
Y	0,857191	0,857438	1,00000	X ₃	-32,52119	-305,5593	76,39721

Dựa vào các bảng kết quả trên. Hãy trả lời các câu hỏi sau với mức ý nghĩa 5%.

- Viết hàm SRF. Giải thích ý nghĩa các hệ số hồi quy.
- Mô hình trên có phù hợp không?
- Tính giá trị nhân tử phóng đại phương sai?
- Mô hình trên có xảy ra hiện tượng đa cộng tuyến không?
- Tìm độ lệch chuẩn của $(\hat{\beta}_2 - \hat{\beta}_3)$.

ĐỀ 7

Khảo sát tiền lương Y (triệu đồng) của một giáo viên theo số năm công tác X (năm) và trình độ học vấn Z ta có bảng số liệu sau :

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
Z	TS	ThS	TS	TS	ThS	ThS	TS	TS	TS	ThS

Câu 1. Giả sử Y và X có quan hệ tuyến tính

- Hãy ước lượng hàm hồi quy của tiền lương theo số năm công tác và nêu ý nghĩa của các hệ số hồi quy tìm được.
- Tính hệ số xác định mô hình và giải thích ý nghĩa.
- Tìm khoảng tin cậy của hệ số hồi quy tổng thể β_2 , với độ tin cậy 95%.
- Dự báo giá trị trung bình của tiền lương khi số năm công tác là 11 năm, với độ tin cậy 95%. Giải thích kết quả.
- Với mức ý nghĩa 5%, hãy cho biết số năm công tác có ảnh hưởng đến tiền lương hay không?
- Ta có kết quả khi dùng kiểm định White như sau :

White Heteroskedasticity Test:			
F-statistic	0.870508	Probability	0.459596
Obs*R-squared	1.991779	Probability	0.369395

Có phương sai thay đổi trong mô hình không, tại sao?

g. Ta có kết quả sau khi dùng kiểm định BG như sau :

Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	0.022578	Probability	0.884798
Obs*R-squared	0.032150	Probability	0.857698

Có tự tương quan bậc nhất trong mô hình không, tại sao?

Câu 2. Với $Z = 0$: Thạc sĩ (ThS); $Z = 1$: Tiến sĩ (TS). Ta có mô hình sau :

$$\hat{Y} = 1,767 + 0,428X + 0,869Z$$

$$se = (0,262) \quad (0,036) \quad (0,209)$$

$$R^2 = 0,957$$

- Nêu ý nghĩa của các hệ số hồi quy riêng?
- Theo bạn thì trình độ học vấn có ảnh hưởng đến tiền lương không? Kết luận với mức ý nghĩa 5%.
- Với mức ý nghĩa 5%, mô hình trên có phù hợp không?
- Ta có kết quả sau khi dùng kiểm định White như sau :

White Heteroskedasticity Test:			
F-statistic	1.025204	Probability	0.475761
Obs*R-squared	4.505987	Probability	0.341838

Có phương sai thay đổi trong mô hình không, tại sao?

e. Ta có kết quả sau khi dùng kiểm định BG như sau :

Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	0.108233	Probability	0.753345
Obs*R-squared	0.177192	Probability	0.673797

Có tự tương quan bậc nhất trong mô hình không, tại sao?

f. Bạn chọn mô hình ở câu 1 hay câu 2, tại sao?

Đề 8

Câu 1. Quan sát về lượng hàng bán được của mặt hàng A (Y-tấn/tháng) và chi phí quảng cáo (X- triệu đồng/tháng) ở một số khu vực bán hàng, người ta thu được các số liệu sau:

X	2	3	4	5	5	6	6	7	7	8
Y	32	32	35	34	36	36	37	38	40	43

a. Hãy ước lượng hàm hồi quy tuyến tính mẫu và nêu ý nghĩa của các hệ số hồi quy tìm được.

- b. Tính hệ số xác định mô hình và giải thích ý nghĩa.
- c. Tìm khoảng tin cậy của các hệ số hồi quy tổng thể, với độ tin cậy 95% và giải thích ý nghĩa.
- d. Tìm khoảng ước lượng cho giá trị trung bình của Y khi chi phí quảng cáo là 9 triệu đồng/tháng, với độ tin cậy 95%. Giải thích kết quả.
- e. Tìm khoảng ước lượng cho giá trị cá biệt của Y khi chi phí quảng cáo là 9 triệu đồng/tháng, với độ tin cậy 95%. Giải thích kết quả.
- f. Theo bạn, chi phí quảng cáo có ảnh hưởng đến lượng hàng bán được hay không? Kết luận với mức ý nghĩa 5%.

Câu 2. Sau đây là hàm hồi quy của doanh số bán (Y : triệu đồng) theo giá bán (X : triệu đồng) và chi phí chào hàng (Z : triệu đồng) với cỡ mẫu là 20 như sau :

$$\begin{aligned} \hat{Y} &= 5,321 - 2,124X + 2,678Z \\ t &= (3,123) \quad (4,221) \quad (5,234) \\ R^2 &= 0,9 \end{aligned}$$

- a. Hãy nêu ý nghĩa của các hệ số hồi quy.
- b. Kiểm định sự phù hợp của mô hình hồi quy, với mức ý nghĩa 1%.
- c. Chi phí chào hàng có ảnh hưởng đến doanh số bán không? Kết luận với mức ý nghĩa 5%.
- d. Tính hệ số xác định mô hình có hiệu chỉnh.

Đề 9

Câu 1. Số liệu về Doanh số bán (Y triệu đồng) và Giá bán (X ngàn đồng/kg) của một loại hàng. Được cho trong bảng sau :

Y	10	9.2	9	8.5	8	7.8	7.3	7	6.5	6.3
X	4.6	5	5.2	6	7.3	7.5	7.9	8	8.5	9

Giả sử Y và X có quan hệ tuyến tính

- a. Hãy ước lượng hàm hồi quy của Doanh số bán phụ thuộc vào Giá bán và nêu ý nghĩa của các hệ số hồi quy tìm được.
- b. Tính hệ số xác định mô hình và giải thích ý nghĩa.
- c. Tìm khoảng tin cậy của các hệ số hồi quy tổng thể, với độ tin cậy 95% và giải thích ý nghĩa.
- d. Dự báo giá trị trung bình của Doanh số bán khi Giá bán là 7 ngàn đồng/kg, với độ tin cậy 95%. Giải thích kết quả.

e. Dự báo giá trị cá biệt của Doanh số bán khi Giá bán là 7 ngàn đồng/kg, với độ tin cậy 95%. Giải thích kết quả.

f. Theo bạn, Giá bán thay đổi có ảnh hưởng đến Doanh số bán? Kết luận với mức ý nghĩa 5%.

Câu 2. Người ta cho rằng mức lương (Y : triệu đồng) không chỉ phụ thuộc vào số năm công tác (X : năm) mà còn phụ thuộc vào trình độ học vấn của người đó ($D = 1$ nếu là tiến sĩ; $D = 0$ nếu là thạc sĩ). Với số liệu của một mẫu có kích thước $n = 10$ người ta đã ước lượng mô hình sau :

$$\hat{Y} = 1,767 + 0,428X + 0,869D$$

$$se = (0,262) \quad (0,036) \quad (0,209)$$

- Hãy nêu ý nghĩa của các hệ số hồi quy.
- Hãy ước lượng các hệ số hồi quy β_2 , với độ tin cậy 95%.
- Theo bạn thì trình độ học vấn có ảnh hưởng đến lương không? Kết luận với mức ý nghĩa 5%.

ĐỀ 10

Câu 1. Bảng sau cho biết số liệu về tổng thu nhập (X : tỷ USD) và mức thuế (Y : tỷ USD).

X	148.23	75.26	181.32	131.88	69.84	67.25	34.82	66.76	7.23	8.07	20.43	33.45
Y	19.74	10.07	26.72	17.82	9.14	9.15	4.5	9.04	0.96	1.13	2.69	4.57

Biết rằng Y và X có quan hệ tuyến tính với nhau

- Hãy ước lượng hàm hồi quy của Y theo X . Giải thích ý nghĩa kinh tế của các hệ số hồi quy nhận được.
- Tính hệ số xác định mô hình và giải thích ý nghĩa của kết quả nhận được.
- Tim khoảng tin cậy cho các hệ số hồi quy tổng thể, với độ tin cậy 95%.
- Với mức ý nghĩa 5%, hãy cho biết khi tổng thu nhập thay đổi có làm ảnh hưởng đến mức thuế không ?
- Với mức tổng thu nhập $X_0 = 150$, hãy dự báo giá trị trung bình và giá trị cá biệt của mức thuế, với độ tin cậy 95%.

Câu 2. Người ta cho rằng chi tiêu cho mặt hàng A (Y : ngàn đồng/tháng) không chỉ phụ thuộc vào thu nhập của người tiêu dùng (X : triệu đồng/tháng) mà còn phụ thuộc vào giới tính của người đó ($D = 1$ nếu là nam; $D = 0$ nếu là nữ). Với số liệu của một mẫu có kích thước $n = 20$ người ta đã ước lượng mô hình sau :

$$\hat{Y} = 0,07 + 0,332D + 0,164X - 0,098XD$$

$$t = (1,947) \quad (6,608) \quad (11,658) \quad (-5,303)$$

- Hãy nêu ý nghĩa của các hệ số hồi quy.
- Hãy ước lượng các hệ số hồi quy, với độ tin cậy 95%.
- Chi tiêu về loại hàng A của nam và nữ có giống nhau hay không? Kết luận với mức ý nghĩa 1%.

Đề 11

Câu 1. Bảng sau cho biết số liệu về tổng thu nhập (X : tỷ USD) và mức thuế (Y : tỷ USD).

X	148.23	75.26	181.32	131.88	69.84	67.25	34.82	66.76	7.23	8.07	20.43	33.45
Y	19.74	10.07	26.72	17.82	9.14	9.15	4.5	9.04	0.96	1.13	2.69	4.57

Giả sử Y và X có quan hệ tuyến tính

- Hãy ước lượng hàm hồi quy của mức thuế phụ thuộc vào tổng thu nhập và nêu ý nghĩa kinh tế của các hệ số hồi quy tìm được.
- Tính hệ số xác định mô hình và giải thích ý nghĩa.
- Tìm khoảng tin cậy của các hệ số hồi quy tổng thể, với độ tin cậy 95% và giải thích ý nghĩa.
- Tìm khoảng tin cậy phương sai của sai số ngẫu nhiên tổng thể, với độ tin cậy 95% .
- Dự báo giá trị trung bình và cá biệt của mức thuế khi tổng thu nhập là 152 tỷ USD, với độ tin cậy 95%. Giải thích kết quả.
- Theo bạn, thu nhập thay đổi có ảnh hưởng đến mức thuế hay không? Kết luận với mức ý nghĩa 5%.

Câu 2. Người ta cho rằng mức lương (Y : triệu đồng) không chỉ phụ thuộc vào số năm công tác (X : năm) mà còn phụ thuộc vào trình độ học vấn của người đó (D = 1 nếu là tiến sĩ; D = 0 nếu là thạc sĩ). Với số liệu của một mẫu có kích thước n = 20 người ta đã ước lượng mô hình sau :

$$\hat{Y} = 1,767 + 0,428X + 0,869D$$

$$se = (0,262) \quad (0,036) \quad (0,209)$$

- Hãy nêu ý nghĩa của các hệ số hồi quy.
- Hãy ước lượng các hệ số hồi quy tổng thể, với độ tin cậy 95%.
- Theo bạn thì trình độ học vấn có ảnh hưởng đến lương không? Kết luận với mức ý nghĩa 5%.

Đề 12

Khảo sát mẫu thống kê như sau :

Y	300	500	400	550	500	600	580	620	600	650
X	1,5	1,8	2	2,1	2,5	2,8	3	3,2	3,4	4
D	Nữ	Nam	Nữ	Nam	Nữ	Nam	Nữ	Nam	Nữ	Nam

Trong đó :

Y là chi tiêu loại hàng A của người tiêu dùng (ngàn đồng/tháng)

X là thu nhập của người tiêu dùng (triệu đồng/tháng)

D là giới tính

Câu 1. Giả sử Y và X có quan hệ tuyến tính

a. Hãy ước lượng hàm hồi quy của chi tiêu theo thu nhập và nêu ý nghĩa của các hệ số hồi quy tìm được.

b. Tính hệ số xác định mô hình và giải thích ý nghĩa.

c. Tìm khoảng tin cậy của hệ số hồi quy tổng thể β_2 , với độ tin cậy 95%.

d. Dự báo giá trị trung bình của chi tiêu khi thu nhập là 2,4 triệu đồng/tháng, với độ tin cậy 95%. Giải thích kết quả.

e. Với mức ý nghĩa 5%, hãy cho biết khi thu nhập thay đổi có ảnh hưởng đến chi tiêu hay không?

f. Tính hệ số co giãn của chi tiêu loại hàng A theo thu nhập tại điểm (\bar{X}, \bar{Y}) và nêu ý nghĩa.

g. Ta có kết quả khi dùng kiểm định White như sau :

White Heteroskedasticity Test:			
F-statistic	13.36609	Probability	0.004071
Obs*R-squared	7.924830	Probability	0.019017

Có hiện tượng phương sai thay đổi trong mô hình trên không? Kết luận với mức ý nghĩa 5%.

h. Ta có kết quả sau khi dùng kiểm định BG như sau :

Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	3.445950	Probability	0.105780
Obs*R-squared	3.298839	Probability	0.069329

Có hiện tượng tự tương quan bậc nhất trong mô hình trên không? Kết luận với mức ý nghĩa 5%.

Câu 2. Đặt $D=0$ nếu là Nữ; $D=1$ nếu là Nam. Ta có mô hình sau :

$$\hat{Y} = 0,07 + 0,332D + 0,164X - 0,098XD$$

$$t \quad (1,947) \quad (6,608) \quad (11,658) \quad (-5,303)$$

$$R^2 = 0,957$$

- Nêu ý nghĩa của các hệ số hồi quy ?
- Tính hệ số xác định mô hình có hiệu chỉnh và cho biết nên chọn mô hình trên hay mô hình câu 1.
- Chi tiêu về loại hàng A của nam và nữ có giống nhau hay không? Kết luận với mức ý nghĩa 5%.
- Hãy ước lượng hệ số hồi quy tổng thể β_3 với độ tin cậy 95%?

Đề 13

Câu 1. Bảng sau cho biết số liệu về tổng thu nhập (X) và tổng chi tiêu cho chăm sóc sức khỏe (Y).

X	21,2	19,3	25,1	27,4	23,3	31,6	31,7	26,4	29,4	30,0	38,8	50,3
Y	3,4	2,3	3,5	3,5	3,4	3,7	4,4	3,9	5,2	4,1	6,1	6,9

Biết rằng Y và X có quan hệ tuyến tính với nhau.

- Hãy ước lượng hàm hồi quy của Y theo X. Giải thích ý nghĩa của các hệ số hồi quy.
 - Tính hệ số xác định mô hình và giải thích ý nghĩa của kết quả nhận được.
 - Tìm khoảng tin cậy cho các hệ số hồi quy với độ tin cậy 95%.
 - Với mức ý nghĩa 5%, hãy cho biết khi X thay đổi có làm ảnh hưởng đến Y không ?
 - Với $X_0 = 38$, hãy dự báo cho giá trị trung bình và giá trị cá biệt của Y, với độ tin cậy 95%.
- Câu 2.** Mối quan hệ giữa giá bán (Y), diện tích (X_2), số phòng tắm (X_3), số phòng ngủ (X_4).

Được thể hiện trong các bảng kết quả dưới đây.

Bảng kết quả hồi quy :

Dependent Variable: Y		Included observations: 14		
	Coefficient	Std. Error	t-Statistic	Prob.
C	129.0616	88.30326	1.461573	0.1746
X2	0.154800	0.031940	4.846516	0.0007
X3	-12.19276	43.25000	-0.281913	0.7838
X4	-21.58752	27.02933	-0.798670	0.4430
R-squared	0.835976	Prob(F-statistic)	0.000299	

Ma trận tương quan :

	X2	X3	X4	Y
X2	1.000000	0.787318	0.464730	0.905827
X3	0.787318	1.000000	0.532327	0.669612
X4	0.464730	0.532327	1.000000	0.315634
Y	0.905827	0.669612	0.315634	1.000000

Dựa vào các bảng kết quả trên. Hãy trả lời câu hỏi sau với mức ý nghĩa $\alpha = 5\%$.

- Viết hàm SRF. Giải thích ý nghĩa của các hệ số hồi quy.
- Mô hình trên có phù hợp với thực tế không ?
- Mô hình trên có xảy ra hiện tượng đa cộng tuyến không ?
- Ước lượng các hệ số hồi quy tổng thể.

Đề 14

Câu 1. Bảng sau cho biết số liệu về doanh thu (X : tỷ USD) và lợi nhuận (Y : triệu USD).

X	37	17	21	30	28	13	26	10	18	12	14	15
Y	629	180	349	453	757	191	490	90	243	168	90	100

Biết rằng Y và X có quan hệ tuyến tính với nhau.

- Hãy ước lượng hàm hồi quy của Y theo X. Giải thích ý nghĩa của các hệ số hồi quy.
- Tính hệ số xác định mô hình và giải thích ý nghĩa của kết quả nhận được.
- Tìm khoảng tin cậy cho các hệ số hồi quy với độ tin cậy 95%.
- Với mức ý nghĩa 5%, hãy cho biết khi X thay đổi có làm ảnh hưởng đến Y không ?
- Với $X_0 = 36$, hãy dự báo cho giá trị trung bình và giá trị cá biệt của Y, với độ tin cậy 95%.

Câu 2. Mối quan hệ giữa tỷ lệ nghèo (Y), số thành viên trong gia đình (X_2), thu nhập bình quân của gia đình (X_3), trình độ học vấn của chủ hộ (X_4) được cho trong các bảng kết quả sau :

Bảng kết quả hồi quy :

Dependent Variable: Y		Included observations: 58		
	Coefficient	Std. Error	t-Statistic	Prob.
C	3.052833	2.972059	1.027178	0.3089
X2	8.484179	1.119004	7.581906	0.0000
X3	-0.630443	0.058575	-10.76298	0.0000
X4	0.335222	0.065866	5.089421	0.0000
R-squared	0.815219	Prob(F-statistic)	0.000000	

Ma trận hiệp phương sai :

	C	X2	X3	X4
C	8.833133	-3.071611	0.018890	-0.062941
X2	-3.071611	1.252169	-0.028417	0.037590
X3	0.018890	-0.028417	0.003431	-0.003388
X4	-0.062941	0.037590	-0.003388	0.004338

Dựa vào các bảng kết quả trên. Hãy trả lời câu hỏi sau với mức ý nghĩa $\alpha = 5\%$.

- Viết hàm SRF. Giải thích ý nghĩa của các hệ số hồi quy.
- Mô hình trên có phù hợp với hay không?
- Hãy kiểm định giả thuyết cho rằng trình độ học vấn và thu nhập của chủ hộ ảnh hưởng như nhau đến tỷ lệ nghèo.
- Ước lượng các hệ số hồi quy tổng thể.

ĐỀ 15

Câu 1. Số liệu về Doanh số bán (Y triệu đồng) và Giá bán (X ngàn đồng/kg) của một loại hàng. Được cho trong bảng sau :

Y	10	9,2	9,0	8,5	8,1	7,8	7,3	7,1	6,5	6,3
X	4,8	5,0	5,2	6,1	7,3	7,5	7,9	8,2	8,5	9,3

Giả sử Y và X có quan hệ tuyến tính

- Hãy ước lượng hàm hồi quy của Doanh số bán phụ thuộc vào Giá bán và nêu ý nghĩa của các hệ số hồi quy tìm được.
- Tính hệ số xác định mô hình và giải thích ý nghĩa.
- Tìm khoảng tin cậy cho các hệ số hồi quy tổng thể, với độ tin cậy 95%.
- Dự báo giá trị trung bình của Doanh số bán khi Giá bán là 10 ngàn đồng/kg, với độ tin cậy 95%. Giải thích kết quả.

Câu 2. Cho kết quả xuất ra từ phần mềm Eview như sau (Số liệu được lấy từ bài tập nhóm lớp 10DNH2)

Dependent Variable: Y				
Method: Least Squares				
Date: 05/12/12 Time: 11:07				
Sample: 1 50				
Included observations: 50				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	8.908000	4.653854	1.914113	0.0620
X2	0.335328	0.056846	5.898921	0.0000
X3	-0.537553	0.235754	-2.280142	0.0274
X4	-0.251578	0.084812	-2.966303	0.0048
X5	0.804397	0.204029	3.942566	0.0003
R-squared	0.967234	Mean dependent var	10.98000	
Adjusted R-squared	0.964321	S.D. dependent var	6.132849	
S.E. of regression	1.158423	Akaike info criterion	3.226636	
Sum squared resid	60.38750	Schwarz criterion	3.417839	
Log likelihood	-75.66591	F-statistic	332.0913	
Durbin-Watson stat	1.537562	Prob(F-statistic)	0.000000	

Trong đó

Y : Số lần sinh viên ăn trưa tại căn tin trong 1 tháng (lần/tháng)

X_2 : Khả năng chi trả (ngàn đồng/bữa)

X_3 : Giá bán thức ăn tại căn tin (ngàn đồng/bữa)

X_4 : Thời gian chờ phục vụ (phút)

X_5 : Số món ăn có trong thực đơn trong 1 ngày (món)

1. Viết hàm SRF.
2. Mô hình trên có phù hợp hay không?
3. Mô hình trên có xảy ra hiện tượng tự tương quan hay không?
4. Giá bán thức ăn tại căn tin có ảnh hưởng đến số lượng sinh viên ăn trưa tại đây hay không?

Câu 3. Cho kết quả các kiểm định sau

Kiểm định 1.

White Heteroskedasticity Test:			
F-statistic	0.931641	Probability	0.428812
Obs*R-squared	2.058254	Probability	0.357319

Kiểm định 2.

Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	0.296972	Probability	0.750914
Obs*R-squared	0.829342	Probability	0.660557

Kiểm định 3.

Dependent Variable: LOG(RESID^2)				
Method: Least Squares				
Date: 05/12/12 Time: 10:57				
Sample: 1 12				
Included observations: 12				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.231418	2.591904	-0.860919	0.4095
LOG(X)	2.442273	1.821531	1.340780	0.2097

Mục đích và kết quả của các kiểm định trên là gì.

ĐỀ 16

Câu 1. Từ một bộ số liệu gồm 20 quan sát, người ta thu được mô hình như sau:

Dependent Variable: Y				
Included observations: 20				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7.480	2.039	3.667	0.000
X	1.218	0.303	4.008	0.002
R-squared	0.693	F-statistic		
Adjusted R-squared	0.480	Prob(F-statistic)		0.008

Trong đó:

Y : Tỷ lệ trẻ em được đi học (%)

X : Thu nhập bình quân của cha mẹ (triệu đồng)

Với mức ý nghĩa thống kê 1%, bạn hãy trả lời các yêu cầu sau đây:

- Viết hàm SRF. Giải thích ý nghĩa kinh tế của các hệ số hồi quy.
- Kiểm định sự ảnh hưởng của thu nhập bình quân của cha mẹ đến tỷ lệ trẻ em tới trường?
- Tìm khoảng ước lượng cho các hệ số hồi quy?
- Kiểm định sự phù hợp của mô hình và cho biết ý nghĩa của hệ số xác định mô hình?
- Có người cho rằng thu nhập bình quân của cha mẹ tăng lên 1 triệu đồng thì tỷ lệ trẻ em được đi học sẽ tăng lên 1,8%. Theo bạn nhận định này có đúng không? Tại sao?

Câu 2. Bảng sau cho biết số liệu về tổng thu nhập (X : tỷ USD) và mức thuế (Y : tỷ USD).

X	148,23	75,26	181,32	131,88	69,84	67,25	34,82	66,76	7,23	8,07	20,43	33,45
Y	19,74	10,07	26,72	17,82	9,14	9,15	4,5	9,04	0,96	1,13	2,69	4,57

Biết rằng Y và X có quan hệ tuyến tính với nhau

Hãy dự báo giá trị trung bình và cá biệt của mức thuế khi tổng thu nhập là 150 tỷ USD với độ tin cậy 95%.

Câu 3. Cho kết quả xuất ra từ phần mềm Eviews như sau (Số liệu được trích từ bài tập nhóm của lớp ĐHTCK6)

Dependent Variable: PRICE				
Method: Least Squares				
Date: 05/10/12 Time: 10:44				
Sample: 1 180				
Included observations: 180				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1298.403	454.9740	-2.853795	0.0048
GTG	76.04801	18.39487	4.134197	0.0001
DTKT	1.437196	0.094313	15.23850	0.0000
SL	-0.000996	0.000460	-2.165496	0.0317
R-squared	0.578833	Mean dependent var	1672.006	
Adjusted R-squared	0.571654	S.D. dependent var	1963.836	
S.E. of regression	1285.294	Akaike info criterion	17.17733	
Sum squared resid	2.91E+08	Schwarz criterion	17.24829	
Log likelihood	-1541.960	F-statistic	80.62887	
Durbin-Watson stat	1.777624	Prob(F-statistic)	0.000000	

Trong đó :

PRICE: Giá thành in bao bì (ngàn đồng/cái)

GTG: Giá thành giấy tại thời điểm in (triệu/tấn)

DTKT : Diện tích khổ trái (cm^2)

SL : Số lượng in (cái)

- Viết hàm SRF. Nêu ý nghĩa các hệ số hồi quy.
- Giải thích ý nghĩa hệ số xác định mô hình và kiểm định sự phù hợp của mô hình.
- Mô hình trên có xảy ra hiện tượng tự tương quan hay không ?

ĐỀ 17

Câu 1. Cho biết mối quan hệ giữa số lượng băng sáng chế (Y : ngàn cái) và chi phí cho nghiên cứu và phát triển (X : tỉ USD) như sau:

Y	90	91	93	95	98	100	102	105	110	115
X	66	70	76	80	84	86	88	90	92	95

Giả sử X và Y có quan hệ tuyến tính với nhau

- Ước lượng hàm hồi quy tuyến tính của Y theo X. Nêu ý nghĩa hệ số góc.
- Tính hệ số xác định mô hình và giải thích kết quả nhận được.
- Tìm khoảng tin cậy cho hệ số hồi quy tổng thể của biến X, với độ tin cậy 95%.

Giải thích kết quả.

- Với mức ý nghĩa 5%, hãy cho biết chi phí cho nghiên cứu và phát triển thay đổi có ảnh hưởng tới số lượng băng sáng chế hay không?

5. Dự báo giá trị trung bình của số lượng bằng sang chế khi chi phí cho nghiên cứu và phát triển là 100 tỉ USD với độ tin cậy 95%. Giải thích kết quả.

Câu 2. Cho hàm hồi quy tuyến tính mẫu như sau

$$\hat{Y} = 1036 - 128X + 163Z - 33Z^2$$

Trong đó :

Y : Số sản phẩm bán được (sản phẩm/tuần)

X : Giá bán (ngàn đồng/sản phẩm)

Z : Chi phí quảng cáo (100000 đồng/tuần)

1. Hãy cho biết dấu của các hệ số hồi quy có giống như những gì ta kỳ vọng hay không ? Giải thích ngắn gọn.
2. Nếu giữ nguyên chi phí quảng cáo và giảm giá 1 ngàn đồng trên sản phẩm thì số sản phẩm bán được tăng trung bình bao nhiêu trong một tuần ?
3. Giả sử giữ nguyên giá ở mức nào đó. Bạn hãy đề xuất mức chi phí quảng cáo tối đa để số sản phẩm bán được là nhiều nhất trong một tuần.

Câu 3. Cho kết quả xuất ra từ phần mềm Eviews như sau

Dependent Variable: BUSTRAVL				
Method: Least Squares				
Date: 04/27/12 Time: 06:54				
Sample: 1 40				
Included observations: 40				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2815.703	976.3007	2.884053	0.0066
INCOME	-0.201273	0.062101	-3.241076	0.0026
POP	1.576575	0.120612	13.07148	0.0000
DENSITY	0.153421	0.034898	4.396311	0.0001
R-squared	0.918759	Mean dependent var	1933.175	
Adjusted R-squared	0.911989	S.D. dependent var	2431.757	
S.E. of regression	721.4228	Akaike info criterion	16.09497	
Sum squared resid	18736228	Schwarz criterion	16.26386	
Log likelihood	-317.8993	F-statistic	135.7080	
Durbin-Watson stat	1.878671	Prob(F-statistic)	0.000000	

Trong đó :

BUSTRAVL : Mức độ giao thông bằng xe buýt ở đô thị (ngàn hành khách mỗi giờ)

INCOME : Thu nhập bình quân đầu người (USD)

POP : Dân số thành phố (ngàn người)

DENSITY : Mật độ dân số (người/dặm vuông)

- Viết hàm SRF. Hãy cho biết dấu của các hệ số hồi quy có giống những gì ta kỳ vọng hay không ?
- Mô hình trên có phù hợp không ?
- Mô hình trên có xảy ra hiện tượng tự tương quan hay không ?

ĐỀ 18

Cho một mẫu gồm các giá trị quan sát sau :

Y_i	10	6	5	8	7	8	7	7	8	9
X_i	2	3	4	2	3	3	4	3	3	2
Z_i	12	9	9	10	9	11	10	9	11	11

Trong đó : Y là lượng khách đi xe buýt (đơn vị : 100 ngàn người).

X là giá vé (đơn vị : ngàn đồng).

Z là giá xăng (đơn vị ngàn đồng/lít).

Câu 1. Giả sử Y và X có quan hệ tuyến tính

- Hãy ước lượng hàm hồi quy của Y theo X. Nêu ý nghĩa kinh tế của các hệ số hồi quy tìm được.
- Kiểm định sự phù hợp của mô hình với mức ý nghĩa 1%.
- Tìm khoảng tin cậy cho các hệ số hồi quy tổng thể, với độ tin cậy 95%.
- Kiểm định giả thuyết $H_0 : \beta_2 = -1,4$ với mức ý nghĩa 5%.
- Dự báo lượng khách đi xe buýt trung bình khi giá vé là 3,5 ngàn đồng, với độ tin cậy 95%.

Câu 2. Với số liệu đã cho, ước lượng mô hình $Y_i = \beta_1 + \beta_2 X_i + \beta_3 Z_i + \varepsilon_i$ bằng EVIEWS ta được:

Dependent Variable: Y

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.544503	2.185040	0.706853	0.5025
X	-0.863874	0.251340	-3.437074	0.0109
Z	0.837696	0.168518	4.970958	0.0016
R-squared	0.916230	Mean dependent var		7.500000
Adjusted R-squared	0.892296	S.D. dependent var		1.433721
S.E. of regression	0.470522	Akaike info criterion		1.573378

Sum squared resid	1.549738	Schwarz criterion	1.664153
Log likelihood	-4.866890	F-statistic	38.28125
Durbin-Watson stat	1.620578	Prob(F-statistic)	0.000170

Hãy trả lời các câu hỏi sau với mức ý nghĩa 5%.

- 1) Viết hàm hồi quy mẫu (SRF) và nêu ý nghĩa các hệ số hồi quy riêng.
- 2) Xét xem giá xăng có ảnh hưởng đến lượng khách đi xe buýt không?
- 3) Tìm khoảng tin cậy cho hệ số hồi quy tổng thể β_2 .
- 4) Trong 2 mô hình ở câu 1 và câu 2 nên chọn mô hình nào? Tại sao?

Câu 3. Để nghiên cứu các yếu tố ảnh hưởng tới lãi suất ngân hàng. Anh (chị) hãy đề xuất mô hình nghiên cứu trong đó có một biến phụ thuộc, phải có ít nhất ba biến độc lập ảnh hưởng đến lãi suất ngân hàng và dự báo dấu của các hệ số hồi quy.

Đề 19

Câu 1. Số liệu về lợi nhuận (Y : tỉ VNĐ) và doanh thu (X : tỉ VNĐ) của một số doanh nghiệp thuộc ngành may mặc ở TP Hồ Chí Minh như sau:

Y	8	9	12	14	20	24	30	36	38	42
X	100	105	110	115	120	125	130	135	140	145

Giả sử Y và X có quan hệ tuyến tính

1. Hãy ước lượng hàm hồi quy của lợi nhuận phụ thuộc vào doanh thu. Nêu ý nghĩa hệ số góc.
2. Tính hệ số xác định mô hình và giải thích ý nghĩa.
3. Tìm khoảng tin cậy cho các hệ số hồi quy tổng thể, với độ tin cậy 95%. Giải thích kết quả nhận được.
4. Tìm khoảng tin cậy cho phương sai của sai số ngẫu nhiên tổng thể, với độ tin cậy 95%.
5. Hãy cho biết doanh thu có thực sự ảnh hưởng tới lợi nhuận hay không? với mức ý nghĩa 1%.
6. Dự báo giá trị trung bình của lợi nhuận khi doanh thu là 150 tỉ VNĐ, với độ tin cậy 95%.

Câu 2. Cho kết quả xuất ra từ phần mềm Eview như sau (Số liệu được lấy từ bài tập nhóm lớp 10DNH2)

Dependent Variable: Y

Included observations: 10

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	641.5907	123.6320		0.0013
X2	1.370831	0.024881		0.0000
X3	1.779807	0.149965		0.0000
R-squared	0.999085	Mean dependent var		6058.520
Adjusted R-squared	0.998824	S.D. dependent var		3890.309
S.E. of regression	133.4361	Akaike info criterion		12.86845
Sum squared resid	124636.4	Schwarz criterion		12.95922
F-statistic	3821.520	Durbin-Watson stat		2.522362
Prob(F-statistic)	0.000000			

Trong đó

Y : Tổng doanh thu du lịch phân theo các khu vực kinh tế (tỉ đồng).

X2 : Doanh thu du lịch phân theo khu vực kinh tế ngoài Nhà nước (tỉ đồng).

X3 : Doanh thu du lịch phân theo khu vực kinh tế có vốn đầu tư nước ngoài (tỉ đồng).

1. Viết hàm SRF. Nêu ý nghĩa các hệ số hồi quy.
2. Giải thích ý nghĩa hệ số xác định mô hình và kiểm định sự phù hợp của mô hình.
3. Có ý kiến cho rằng khi doanh thu du lịch phân theo khu vực kinh tế có vốn đầu tư nước ngoài tăng 1 tỉ đồng thì tổng doanh thu du lịch phân theo các khu vực kinh tế tăng hơn 2 tỉ đồng có đúng không? với mức ý nghĩa 5%.

Câu 3. Cho kết quả các kiểm định sau

Kiểm định 1.

Heteroskedasticity Test: White

F-statistic	1.046211	Prob. F(5,4)	0.4964
Obs*R-squared	5.666801	Prob. Chi-Square(5)	0.3400

Scaled explained SS 2.583828 Prob. Chi-Square(5) 0.7638

Kiểm định 2.

Breusch-Godfrey Serial Correlation LM Test:

F-statistic 1.719734 Prob. F(2,5) 0.2702

Obs*R-squared 4.075456 Prob. Chi-Square(2) 0.2303

Kiểm định 3.

Heteroskedasticity Test: Glejser

F-statistic 0.238216 Prob. F(2,7) 0.7942

Obs*R-squared 0.637244 Prob. Chi-Square(2) 0.7272

Scaled explained SS 0.525739 Prob. Chi-Square(2) 0.7688

Mục đích và kết quả của các kiểm định trên là gì

ĐỀ 20

Câu 1. Số liệu về doanh số bán hàng (Y : triệu/tháng) và chi phí quảng cáo (X : triệu/tháng) của một số doanh nghiệp như sau:

Y	1268	1491	1589	1628	1019	1802	1607	1284	1389	1440
X	181	247	222	138	152	261	258	161	172	228

Giả sử Y và X có quan hệ tuyến tính

- Hãy ước lượng hàm hồi quy của doanh số bán hàng phụ thuộc vào chi phí quảng cáo. Nêu ý nghĩa hệ số góc.
- Nếu chi phí quảng cáo tăng 1% thì doanh số bán hàng thay đổi như thế nào?
- Tìm khoảng tin cậy cho các hệ số hồi quy tổng thể, với độ tin cậy 90%.
- Hãy cho biết chi phí quảng cáo có thực sự ảnh hưởng tới doanh số bán hàng hay không? với mức ý nghĩa 5%.
- Dự báo giá trị doanh số bán hàng trung bình khi chi phí quảng cáo là 140 triệu/tháng, với độ tin cậy 95%.
- Hãy cho biết ý kiến sau đúng hay sai với mức ý nghĩa 1% “nếu chi phí quảng cáo tăng 1 triệu/tháng thì doanh số bán hàng tăng hơn 2 triệu/tháng”.

Câu 2. Cho kết quả xuất ra từ phần mềm Eview như sau (Số liệu được lấy từ bài tập nhóm lớp 11DNH4)

Dependent Variable: Y

Included observations: 70

Variable	Coefficien			
	t	Std. Error	t-Statistic	Prob.
C	6.157569	0.224700	27.40347	0.0000
X2	-0.227033	0.044851	-5.061935	0.0000
X3	0.427811	0.055463	7.713495	0.0000
X4	-0.112598	0.033082	-3.403591	0.0011
R-squared	0.929594	Mean dependent var	6.968429	
Adjusted R-squared	0.926394	S.D. dependent var	0.958035	
S.E. of regression	0.259919	Akaike info criterion	0.198553	
F-statistic	290.4741	Durbin-Watson stat	2.044807	
Prob(F-statistic)	0.000000			

Trong đó

Y : Điểm trung bình (Điểm).

X2 : Số ngày nghỉ học/tuần (ngày).

X3 : Số giờ tự học/ngày (giờ)

X4 : Số giờ đi làm thêm/ngày (giờ).

Ma trận tương quan

	Y	X2	X3	X4
Y	1.000000	-0.823130	0.941878	-0.875080
X2	-0.823130	1.000000	-0.752638	0.672805
X3	0.941878	-0.752638	1.000000	-0.863144
X4	-0.875080	0.672805	-0.863144	1.000000

1. Viết hàm SRF. Nêu ý nghĩa các hệ số hồi quy.
2. Giải thích ý nghĩa hệ số xác định mô hình và kiểm định sự phù hợp của mô hình.
3. Mô hình trên có xảy ra hiện tượng tự tương quan hay không?
4. Hãy nêu 4 cách phát hiện đa cộng tuyến. Mô hình trên có xảy ra hiện tượng đa cộng tuyến hay không?

5. Hãy cho biết mục đích và kết quả các kiểm định sau

Kiểm định 1.

Wald Test:

Equation: Untitled

Test Statistic	Value	df	Probability
t-statistic	0.326671	65	0.7450
F-statistic	0.106714	(1, 65)	0.7450
Chi-square	0.106714	1	0.7439

Null Hypothesis: C(4)=0

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(4)	0.016167	0.049491

Restrictions are linear in coefficients.

Kiểm định 2.

Heteroskedasticity Test: White

F-statistic	1.444954	Prob. F(9,60)	0.1898
Obs*R-squared	12.46937	Prob. Chi-Square(9)	0.1881
Scaled explained SS	8.116879	Prob. Chi-Square(9)	0.5224

Đề 21

Câu 1. Cho biết mối quan hệ về lượng cam bán được (Y : tấn/tháng) và giá cam (X : ngàn đồng/kg) như sau:

Y	14	13	12	11	10	9	8	8	6	5	4	3
X	20	25	30	35	36	38	40	43	45	50	55	60

Giả sử Y và X có quan hệ tuyến tính

- Hãy ước lượng hàm hồi quy của lượng cam bán được phụ thuộc vào giá cam. Nêu ý nghĩa hệ số góc.
- Tính hệ số xác định mô hình và giải thích ý nghĩa.

3. Tìm khoảng tin cậy cho các hệ số hồi quy tổng thể, với độ tin cậy 90%.
4. Hãy cho biết giá cam có thực sự ảnh hưởng tới lượng cam bán được hay không? với mức ý nghĩa 5%.
5. Dự báo giá trị lượng cam bán được trung bình khi giá bán là 70 ngàn đồng/kg, với độ tin cậy 95%.
6. Với mức ý nghĩa 5%, nếu giá cam tăng lên 1 ngàn đồng/kg thì lượng cam bán được giảm tối thiểu bao nhiêu?

Câu 2. Cho kết quả xuất ra từ phần mềm Eview như sau (Số liệu được lấy từ bài tập nhóm lớp 10DTC)

Dependent Variable: Y

Included observations: 12

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	206263.2	27871.24	7.400574	0.0001
X2	0.077423	0.317303	0.244003	0.8134
X3	1.042863	0.305588	3.412638	0.0092
X4	-12717.98	4673.433	-2.721336	0.0262
R-squared	0.995294	Mean dependent var	1117000.	
Adjusted R-squared	0.993529	S.D. dependent var	667212.0	
S.E. of regression	53673.80	Akaike info criterion	24.88044	
F-statistic	563.9305	Durbin-Watson stat	1.660274	
Prob(F-statistic)	0.000000			

Trong đó

Y : Tổng sản phẩm quốc nội (tỷ đồng)

X₂: Giá trị xuất khẩu hàng hóa và dịch vụ (tỷ đồng)

X₃: Giá trị nhập khẩu hàng hóa và dịch vụ (tỷ đồng)

X₄: Tỷ lệ lạm phát (%)

Ma trận tương quan

	Y	X2	X3	X4
Y	1.000000	0.993728	0.994929	0.766737
X2	0.993728	1.000000	0.995810	0.790547
X3	0.994929	0.995810	1.000000	0.813496
X4	0.766737	0.790547	0.813496	1.000000

1. Viết hàm SRF. Hãy cho biết dấu của các hệ số hồi quy có giống những gì ta kỳ vọng hay không?
2. Giải thích ý nghĩa hệ số xác định mô hình và kiểm định sự phù hợp của mô hình.
3. Đa cộng tuyến là gì? Mô hình trên có xảy ra hiện tượng đa cộng tuyến không?
4. Cho kiểm định sau

Wald Test:

Equation: Untitled

Test Statistic	Value	df	Probability
t-statistic	0.244003	8	0.8134
F-statistic	0.059537	(1, 8)	0.8134
Chi-square	0.059537	1	0.8072

Null Hypothesis: C(2)=0

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(2)	0.077423	0.317303

Có nên loại X2 ra khỏi mô hình hay không?

ĐỀ 22

Câu 1. Cho biết mối quan hệ về lượng hàng bán được (Y : tấn/tháng) và giá bán (X : ngàn đồng/kg) như sau:

Y	1400	750	1500	1100	1600	850	1350	1550	1420	1200	950	1300
X	41	48	40	45	38	47	43	39	42	44	46	43

Giả sử Y và X có quan hệ tuyến tính

1. Hãy ước lượng hàm hồi quy của lượng hàng bán được phụ thuộc vào giá bán. Nêu ý nghĩa hệ số góc.
2. Tính hệ số xác định mô hình và giải thích ý nghĩa.
3. Tìm khoảng tin cậy cho các hệ số hồi quy tổng thể, với độ tin cậy 90%.
4. Hãy cho biết giá bán có thực sự ảnh hưởng tới lượng hàng bán được hay không? với mức ý nghĩa 5%.
5. Dự báo giá trị lượng hàng bán được trung bình khi giá bán là 50 ngàn đồng/kg, với độ tin cậy 95%.

Câu 2. Lượng hàng bán được không những phụ thuộc vào giá bán mà còn phụ thuộc vào địa điểm bán ($Z = 0$: Nông thôn, $Z = 1$: Thành phố). Ta có kết quả xuất ra từ Eview như sau

Dependent Variable: Y

Included observations: 12

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4832.452	236.6773	20.41790	0.0000
X	-84.14013	5.395466	-15.59460	0.0000
Z	66.14650	32.67117	2.024614	0.0736
R-squared	0.968254	Mean dependent var	1247.500	
Adjusted R-squared	0.961200	S.D. dependent var	280.2312	
F-statistic	137.2521	Durbin-Watson stat	2.338100	
Prob(F-statistic)	0.000000			

1. Viết hàm SRF. Nêu ý nghĩa các hệ số hồi quy.
2. Giải thích ý nghĩa hệ số xác định mô hình và kiểm định sự phù hợp của mô hình.
3. Dự báo (điểm) lượng hàng trung bình bán được tại thành phố khi giá bán là 55 ngàn đồng/kg
4. Mô hình có xảy ra hiện tượng tự tương quan hay không.

Câu 3. Cho kết quả các kiểm định sau

Kiểm định 1.

Heteroskedasticity Test: White

F-statistic	1.046211	Prob. F(5,4)	0.4964
Obs*R-squared	5.666801	Prob. Chi-Square(5)	0.3400
Scaled explained SS	2.583828	Prob. Chi-Square(5)	0.7638

Kiểm định 2.

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.273055	Prob. F(2,7)	0.7688
Obs*R-squared	0.868437	Prob. Chi-Square(2)	0.6478

Kiểm định 3.

Heteroskedasticity Test: Glejser

F-statistic	0.238216	Prob. F(2,7)	0.7942
Obs*R-squared	0.637244	Prob. Chi-Square(2)	0.7272
Scaled explained SS	0.525739	Prob. Chi-Square(2)	0.7688

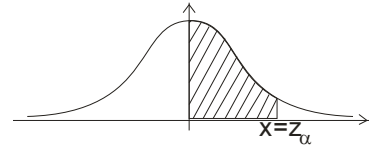
Mục đích và kết quả của các kiểm định trên là gì

PHỤ LỤC

PHÂN PHỐI GAUSS

$$\varphi(x) = \Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt = P(0 \leq X \leq x) \equiv \alpha,$$

với $X \sim N(0;1)$, $x \equiv z_\alpha$.



	0	0.01	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

PHÂN PHỐI STUDENT

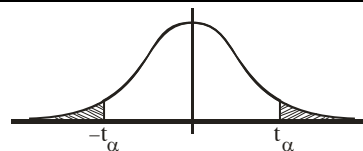
$$P(|T| \geq t_\alpha) = \alpha \text{ với } T \sim St(n)$$

Cột 1 : giá trị độ tự do n.

Hàng 1 : Giá trị nguy cơ sai lầm α

Hàng 2 : Giá trị nguy cơ sai lầm $\alpha / 2$

Nội dung bảng : Giá trị t_α tương ứng với n và α



α	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1	0.15	0.2
$\alpha / 2$	0.005	0.01	0.015	0.02	0.025	0.03	0.035	0.04	0.045	0.05	0.075	0.1
1	63.656	31.821	21.205	15.894	12.706	10.579	9.058	7.916	7.026	6.314	4.165	3.078
2	9.925	6.965	5.643	4.849	4.303	3.896	3.578	3.320	3.104	2.920	2.282	1.886
3	5.841	4.541	3.896	3.482	3.182	2.951	2.763	2.605	2.471	2.353	1.924	1.638
4	4.604	3.747	3.298	2.999	2.776	2.601	2.456	2.333	2.226	2.132	1.778	1.533
5	4.032	3.365	3.003	2.757	2.571	2.422	2.297	2.191	2.098	2.015	1.699	1.476
6	3.707	3.143	2.829	2.612	2.447	2.313	2.201	2.104	2.019	1.943	1.650	1.440
7	3.499	2.998	2.715	2.517	2.365	2.241	2.136	2.046	1.966	1.895	1.617	1.415
8	3.355	2.896	2.634	2.449	2.306	2.189	2.090	2.004	1.928	1.860	1.592	1.397
9	3.250	2.821	2.574	2.398	2.262	2.150	2.055	1.973	1.899	1.833	1.574	1.383
10	3.169	2.764	2.527	2.359	2.228	2.120	2.028	1.948	1.877	1.812	1.559	1.372
11	3.106	2.718	2.491	2.328	2.201	2.096	2.007	1.928	1.859	1.796	1.548	1.363
12	3.055	2.681	2.461	2.303	2.179	2.076	1.989	1.912	1.844	1.782	1.538	1.356
13	3.012	2.650	2.436	2.282	2.160	2.060	1.974	1.899	1.832	1.771	1.530	1.350
14	2.977	2.624	2.415	2.264	2.145	2.046	1.962	1.887	1.821	1.761	1.523	1.345
15	2.947	2.602	2.397	2.249	2.131	2.034	1.951	1.878	1.812	1.753	1.517	1.341
16	2.921	2.583	2.382	2.235	2.120	2.024	1.942	1.869	1.805	1.746	1.512	1.337
17	2.898	2.567	2.368	2.224	2.110	2.015	1.934	1.862	1.798	1.740	1.508	1.333
18	2.878	2.552	2.356	2.214	2.101	2.007	1.926	1.855	1.792	1.734	1.504	1.330
19	2.861	2.539	2.346	2.205	2.093	2.000	1.920	1.850	1.786	1.729	1.500	1.328
20	2.845	2.528	2.336	2.197	2.086	1.994	1.914	1.844	1.782	1.725	1.497	1.325
21	2.831	2.518	2.328	2.189	2.080	1.988	1.909	1.840	1.777	1.721	1.494	1.323
22	2.819	2.508	2.320	2.183	2.074	1.983	1.905	1.835	1.773	1.717	1.492	1.321
23	2.807	2.500	2.313	2.177	2.069	1.978	1.900	1.832	1.770	1.714	1.489	1.319
24	2.797	2.492	2.307	2.172	2.064	1.974	1.896	1.828	1.767	1.711	1.487	1.318
25	2.787	2.485	2.301	2.167	2.060	1.970	1.893	1.825	1.764	1.708	1.485	1.316
26	2.779	2.479	2.296	2.162	2.056	1.967	1.890	1.822	1.761	1.706	1.483	1.315
27	2.771	2.473	2.291	2.158	2.052	1.963	1.887	1.819	1.758	1.703	1.482	1.314
28	2.763	2.467	2.286	2.154	2.048	1.960	1.884	1.817	1.756	1.701	1.480	1.313
29	2.756	2.462	2.282	2.150	2.045	1.957	1.881	1.814	1.754	1.699	1.479	1.311
∞	2.576	2.326	2.170	2.054	1.960	1.881	1.812	1.751	1.695	1.645	1.440	1.282

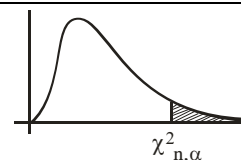
PHÂN PHỐI CHI – BÌNH PHƯƠNG

$$P(X \geq \chi_{n,\alpha}^2) = \alpha \text{ khi } X \sim \chi^2(n)$$

Hàng 1 : Giá trị của α

Cột 1 : Giá trị độ tự do n.

Nội dung bảng : Giá trị $\chi_{n,\alpha}^2$.



	0.005	0.01	0.015	0.02	0.025	0.03	0.05	0.95	0.975	0.98	0.99	0.995
1	7.879	6.635	5.916	5.412	5.024	4.709	3.841	0.004	0.001	0.001	0.000	0.000
2	10.597	9.210	8.399	7.824	7.378	7.013	5.991	0.103	0.051	0.040	0.020	0.010
3	12.838	11.345	10.465	9.837	9.348	8.947	7.815	0.352	0.216	0.185	0.115	0.072
4	14.860	13.277	12.339	11.668	11.143	10.712	9.488	0.711	0.484	0.429	0.297	0.207
5	16.750	15.086	14.098	13.388	12.832	12.375	11.070	1.145	0.831	0.752	0.554	0.412
6	18.548	16.812	15.777	15.033	14.449	13.968	12.592	1.635	1.237	1.134	0.872	0.676
7	20.278	18.475	17.398	16.622	16.013	15.509	14.067	2.167	1.690	1.564	1.239	0.989
8	21.955	20.090	18.974	18.168	17.535	17.011	15.507	2.733	2.180	2.032	1.647	1.344
9	23.589	21.666	20.512	19.679	19.023	18.480	16.919	3.325	2.700	2.532	2.088	1.735
10	25.188	23.209	22.021	21.161	20.483	19.922	18.307	3.940	3.247	3.059	2.558	2.156
11	26.757	24.725	23.503	22.618	21.920	21.342	19.675	4.575	3.816	3.609	3.053	2.603
12	28.300	26.217	24.963	24.054	23.337	22.742	21.026	5.226	4.404	4.178	3.571	3.074
13	29.819	27.688	26.403	25.471	24.736	24.125	22.362	5.892	5.009	4.765	4.107	3.565
14	31.319	29.141	27.827	26.873	26.119	25.493	23.685	6.571	5.629	5.368	4.660	4.075
15	32.801	30.578	29.235	28.259	27.488	26.848	24.996	7.261	6.262	5.985	5.229	4.601
16	34.267	32.000	30.629	29.633	28.845	28.191	26.296	7.962	6.908	6.614	5.812	5.142
17	35.718	33.409	32.011	30.995	30.191	29.523	27.587	8.672	7.564	7.255	6.408	5.697
18	37.156	34.805	33.382	32.346	31.526	30.845	28.869	9.390	8.231	7.906	7.015	6.265
19	38.582	36.191	34.742	33.687	32.852	32.158	30.144	10.117	8.907	8.567	7.633	6.844
20	39.997	37.566	36.093	35.020	34.170	33.462	31.410	10.851	9.591	9.237	8.260	7.434
21	41.401	38.932	37.434	36.343	35.479	34.759	32.671	11.591	10.283	9.915	8.897	8.034
22	42.796	40.289	38.768	37.659	36.781	36.049	33.924	12.338	10.982	10.600	9.542	8.643
23	44.181	41.638	40.094	38.968	38.076	37.332	35.172	13.091	11.689	11.293	10.196	9.260
24	45.558	42.980	41.413	40.270	39.364	38.609	36.415	13.848	12.401	11.992	10.856	9.886
25	46.928	44.314	42.725	41.566	40.646	39.880	37.652	14.611	13.120	12.697	11.524	10.520
26	48.290	45.642	44.031	42.856	41.923	41.146	38.885	15.379	13.844	13.409	12.198	11.160
27	49.645	46.963	45.331	44.140	43.195	42.407	40.113	16.151	14.573	14.125	12.878	11.808
28	50.994	48.278	46.626	45.419	44.461	43.662	41.337	16.928	15.308	14.847	13.565	12.461
29	52.335	49.588	47.915	46.693	45.722	44.913	42.557	17.708	16.047	15.574	14.256	13.121
30	53.672	50.892	49.199	47.962	46.979	46.160	43.773	18.493	16.791	16.306	14.953	13.787
35	60.275	57.342	55.553	54.244	53.203	52.335	49.802	22.465	20.569	20.027	18.509	17.192
40	66.766	63.691	61.812	60.436	59.342	58.428	55.758	26.509	24.433	23.838	22.164	20.707
45	73.166	69.957	67.994	66.555	65.410	64.454	61.656	30.612	28.366	27.720	25.901	24.311
50	79.490	76.154	74.111	72.613	71.420	70.423	67.505	34.764	32.357	31.664	29.707	27.991
55	85.749	82.292	80.173	78.619	77.380	76.345	73.311	38.958	36.398	35.659	33.571	31.735
60	91.952	88.379	86.188	84.580	83.298	82.225	79.082	43.188	40.482	39.699	37.485	35.534
65	98.105	94.422	92.161	90.501	89.177	88.069	84.821	47.450	44.603	43.779	41.444	39.383
70	104.215	100.425	98.098	96.387	95.023	93.881	90.531	51.739	48.758	47.893	45.442	43.275
75	110.285	106.393	104.001	102.243	100.839	99.665	96.217	56.054	52.942	52.039	49.475	47.206
80	116.321	112.329	109.874	108.069	106.629	105.422	101.879	60.391	57.153	56.213	53.540	51.172
85	122.324	118.236	115.720	113.871	112.393	111.156	107.522	64.749	61.389	60.412	57.634	55.170
90	128.299	124.116	121.542	119.648	118.136	116.869	113.145	69.126	65.647	64.635	61.754	59.196
95	134.247	129.973	127.341	125.405	123.858	122.562	118.752	73.520	69.925	68.879	65.898	63.250
100	140.170	135.807	133.120	131.142	129.561	128.237	124.342	77.929	74.222	73.142	70.065	67.328

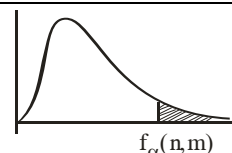
PHÂN PHỐI FISHER

$$P(X \geq f_\alpha(n, m)) = \alpha \text{ khi } X \sim F(n, m)$$

Hàng 1 : Giá trị của độ tự do (tử số) n. **Cột 1 :**

Giá trị độ tự do (mẫu số) m.

Nội dung bảng : Giá trị $f_\alpha(n, m)$.



Bảng 1 : $\alpha = 0.05$

	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253
2	18.51	19	19.16	19.25	19.3	19.33	19.35	19.37	19.38	19.4	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.7	8.66	8.64	8.62	8.59	8.57	8.55
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6	5.96	5.91	5.86	5.8	5.77	5.75	5.72	5.69	5.66
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.5	4.46	4.43	4.4
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.1	4.06	4	3.94	3.87	3.84	3.81	3.77	3.74	3.7
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.3	3.27
8	5.32	4.46	4.07	3.84	3.69	3.58	3.5	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.9	2.86	2.83	2.79	2.75
10	4.96	4.1	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.7	2.66	2.62	2.58
11	4.84	3.98	3.59	3.36	3.2	3.09	3.01	2.95	2.9	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45
12	4.75	3.89	3.49	3.26	3.11	3	2.91	2.85	2.8	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.6	2.53	2.46	2.42	2.38	2.34	2.3	2.25
14	4.6	3.74	3.34	3.11	2.96	2.85	2.76	2.7	2.65	2.6	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18
15	4.54	3.68	3.29	3.06	2.9	2.79	2.71	2.64	2.59	2.54	2.48	2.4	2.33	2.29	2.25	2.2	2.16	2.11
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06
17	4.45	3.59	3.2	2.96	2.81	2.7	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.1	2.06	2.01
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97
19	4.38	3.52	3.13	2.9	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93
20	4.35	3.49	3.1	2.87	2.71	2.6	2.51	2.45	2.39	2.35	2.28	2.2	2.12	2.08	2.04	1.99	1.95	1.9
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.1	2.05	2.01	1.96	1.92	1.87
22	4.3	3.44	3.05	2.82	2.66	2.55	2.46	2.4	2.34	2.3	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84
23	4.28	3.42	3.03	2.8	2.64	2.53	2.44	2.37	2.32	2.27	2.2	2.13	2.05	2.01	1.96	1.91	1.86	1.81
24	4.26	3.4	3.01	2.78	2.62	2.51	2.42	2.36	2.3	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79
25	4.24	3.39	2.99	2.76	2.6	2.49	2.4	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2	1.92	1.84	1.79	1.74	1.69	1.64	1.58
60	4	3.15	2.76	2.53	2.37	2.25	2.17	2.1	2.04	1.99	1.92	1.84	1.75	1.7	1.65	1.59	1.53	1.47
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.5	1.43	1.35
∞	3.84	3	2.6	2.37	2.21	2.1	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22

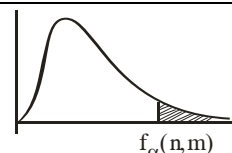
PHÂN PHỐI FISHER

$$P(X \geq f_\alpha(n, m)) = \alpha \text{ khi } X \sim F(n, m)$$

Hàng 1 : Giá trị của độ tự do (tử số) n. **Cột 1 :**

Giá trị độ tự do (mẫu số) m.

Nội dung bảng : Giá trị $f_\alpha(n, m)$.



Bảng 2 : $\alpha = 0.01$

	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120
1	4052	4999	5404	5624	5764	5859	5928	5981	6022	6056	6107	6157	6209	6234	6260	6286	6313	6340
2	98.50	99.00	99.16	99.25	99.30	99.33	99.36	99.38	99.39	99.40	99.42	99.43	99.45	99.46	99.47	99.48	99.48	99.49
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.34	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84
17	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32