

BÀI TẬP KINH TẾ LƯỢNG

(Dành cho sinh viên Hệ chính quy – ĐHNL-2013 tham khảo)

PHẦN CƠ BẢN

Bài 1:

Số liệu về tỷ lệ lạm phát và lãi suất trong năm 2011 của 6 quốc gia như sau

Lãi suất (Y) %	7	11	20	10	16	14
Lạm phát (X) %	3	8	17	8	12	12

Yêu cầu:

- Ước lượng và viết phương trình hồi quy tuyến tính: $Y = \beta_1 + \beta_2 X + \varepsilon_i$
- Tìm hệ số xác định R^2 ? Ý nghĩa?

Bài 2:

Kết quả hồi quy tuyến tính về các yếu tố ảnh hưởng đến điểm trung bình học tập ở học kỳ 1 niên khoá 2008-2009 của 74 sinh viên được cho như sau:

Dependent Variable: DIEM
 Method: Least Squares
 Date: 11/26/09 Time: 15:30
 Sample: 1 74
 Included observations: 74

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.729045	?	29.51623	0.0000
KCACH	-0.010191	0.015832	-0.643724	0.5219
TGTH	0.186606	0.057092	3.268493	0.0017
TGTV	?	0.061147	4.670822	0.0000
DILAM	0.254260	0.144414	1.760624	0.0828
MAYTINH	-0.157488	0.119983	?	0.1937
R-squared	0.528789	Mean dependent var		6.422703
Adjusted R-squared	?	S.D. dependent var		0.698859
S.E. of regression	0.497054	Akaike info criterion		1.517370
Sum squared resid	16.80029	Schwarz criterion		1.704186
Log likelihood	-50.14270	F-statistic		?
Durbin-Watson stat	1.704267	Prob(F-statistic)		0.000000

Mô tả các biến:

- DIEM : Điểm trung bình học kỳ 1 niên khoá 2008-2009
- KCACH : Khoảng cách đi học mỗi ngày (km)
- TGTH : Thời gian tự học (giờ/ngày)
- TGTV : Thời gian học và đọc sách ở thư viện (giờ/ngày)
- DILAM : 1: Có đi làm thêm; 0: Không có đi làm thêm
- MAYTINH : 1: Có máy vi tính; 0: Không có máy vi tính

Yêu cầu:

- (1) Hãy điền số vào các dấu chấm hỏi?
- (2) Viết phương trình hồi quy tuyến tính?
- (3) Giải thích ý nghĩa của hệ số xác định (R -squared)?

- (4) Với độ tin cậy 90%, biến KCACH có ảnh hưởng đến điểm học tập không? Vì sao?
- (5) Với độ tin cậy 90%, mô hình này phù hợp không? Vì sao?
- (6) Giải thích ý nghĩa của biến TGTV trong mô hình?
- (7) Giải thích ý nghĩa của biến DILAM trong mô hình?
- (8) Có người cho rằng sinh viên có máy tính sẽ mê chơi game, nghe nhạc hoặc xem phim nên ảnh hưởng xấu đến kết quả học tập. Theo bạn, với độ tin cậy 90%, bạn có đồng ý với nhận định đó hay không (nếu như sử dụng kết quả mô hình trên để giải thích)?
- (9) Theo bạn, để nghiên cứu các yếu tố ảnh hưởng đến điểm học tập của sinh viên, ta cần đưa vào những biến độc lập nào nữa (nếu có thể)? Giải thích.

Bài 3:

Kết quả hồi quy tuyến tính về các yếu tố ảnh hưởng đến tiền lương của 35 công nhân được cho như sau:

Dependent Variable: WAGE
 Method: Least Squares
 Date: 11/26/09 Time: 16:02
 Sample: 1 35
 Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	767.5716	513.6939	1.494220	0.1452
AGE	-3.624412	10.29980	-0.351892	0.7273
EDUC	124.2512	44.01766	2.822759	0.0082
EXPER	45.82795	17.20317	2.663925	0.0121
R-squared	0.271605	Mean dependent var		1775.371
Adjusted R-squared	0.201115	S.D. dependent var		606.0181
S.E. of regression	541.6611	Akaike info criterion		15.53437
Sum squared resid	9095298.	Schwarz criterion		15.71212
Log likelihood	-267.8514	F-statistic		3.853114
Durbin-Watson stat	1.509535	Prob(F-statistic)		0.018767

Mô tả các biến:

- WAGE : Tiền lương hàng tháng của công nhân (USD/tháng)
- AGE : Tuổi của công nhân (tuổi)
- EDUC : Số lớp đã đi học (năm)
- EXPER : Số năm kinh nghiệm làm việc (năm)

Yêu cầu:

- (1) Viết phương trình hồi quy tuyến tính?
- (2) Giải thích ý nghĩa của hệ số xác định (R-squared)?
- (3) Với độ tin cậy 95%, biến AGE có ảnh hưởng đến WAGE không? Vì sao?
- (4) Với độ tin cậy 95%, mô hình này phù hợp không? Vì sao?
- (5) Giải thích ý nghĩa của biến EDUC trong mô hình?
- (6) Hãy dự báo cho lương của công nhân nếu AGE= 40(tuổi); EDUC= 8(năm) và EXPER= 15(năm)?
- (7) Có người cho rằng công nhân có nhiều năm kinh nghiệm thì sẽ nhận được mức lương cao hơn. Theo bạn, với độ tin cậy 95%, bạn có đồng ý với nhận định đó hay không (nếu như sử dụng kết quả mô hình trên để giải thích)?
- (8) Theo bạn, để nghiên cứu các yếu tố ảnh hưởng đến WAGE, ta có thể đưa vào những biến độc lập nào nữa (nếu có thể)? Giải thích.

Bài 4:

Kết quả hồi quy tuyến tính về nhu cầu tiêu dùng thuốc lá tại Thổ Nhĩ Kỳ giai đoạn 1960-1988 (n=29), Nguồn: Aysit Tansel, 1993, "Cigarette Demand, Health Scares, and Education in Turkey" trong quyển *Applied Economics*, trang 521-529, như sau:

Dependent Variable: Q
 Method: Least Squares
 Date: 10/15/09 Time: 19:43
 Sample: 1960 1988
 Included observations: 29

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.927171	0.366994	2.526390	0.0182
GNP	0.000804	0.000226	3.559658	0.0015
PRICE	-0.328197	0.103130	-3.182371	0.0039
EDU	-0.049489	0.023594	-2.097542	0.0462
R-squared	0.684650	Mean dependent var		2.204655
Adjusted R-squared	?	S.D. dependent var		0.243190
S.E. of regression	0.144527	Akaike info criterion		-0.903252
Sum squared resid	0.522205	Schwarz criterion		-0.714660
Log likelihood	17.09716	F-statistic		?
Durbin-Watson stat	0.980035	Prob(F-statistic)		0.000002

Mô tả các biến:

- Q : Tiêu dùng thuốc lá tính trên đầu người (kg)
- GNP : GNP bình quân đầu người (đồng Lira Thổ Nhĩ Kỳ)
- PRICE : Giá 1 kg thuốc lá (đồng Lira Thổ Nhĩ Kỳ)
- EDU : Tỷ lệ đăng ký học cấp 2 và cấp 3 trong độ tuổi 12-17 (%)

Yêu cầu:

- (1) Viết phương trình hồi quy tuyến tính?
- (2) Giải thích ý nghĩa của hệ số xác định (R-squared)?
- (3) Với độ tin cậy 95%, mô hình này phù hợp không? Vì sao?
- (4) Tìm Adjusted R-squared, thống kê F để thay vào dấu chấm hỏi của bảng kết xuất?
- (5) Viết công thức t bảng, biết $\alpha=5\%$, kiểm định 2 đuôi?
- (6) Viết công thức F bảng, biết $\alpha=5\%$?
- (7) Giải thích ý nghĩa của biến GNP trong mô hình?
- (8) Số liệu dùng để ước lượng hồi quy thuộc kiểu dữ liệu nào? Kiểu này thường vi phạm giả thiết gì khi xây dựng phương trình hồi quy tuyến tính (OLS)?
- (9) Có người cho rằng với độ tin cậy 95%, ở giai đoạn 1960-1988, giá thuốc lá không ảnh hưởng đến nhu cầu tiêu thụ thuốc lá của người dân Thổ Nhĩ Kỳ. Quan điểm của bạn như thế nào (nếu sử dụng kết quả mô hình để giải thích)?
- (10) Ý nghĩa của cột Prob?

Giải bài 4

- (1) Viết phương trình hồi quy tuyến tính:

$$Q = 0.9271 + 0.0008 \cdot \text{GNP} - 0.3282 \cdot \text{P} - 0.0495 \cdot \text{EDU}$$
- (2) Giải thích ý nghĩa của hệ số xác định (R-squared):
 R-squared = $R^2 = 0.684650$. Mô hình hồi quy này giải thích được 68% sự biến thiên của nhu cầu tiêu thụ thuốc lá của Thổ Nhĩ Kỳ (giai đoạn 1960-1988)
- (3) Với độ tin cậy 95%, mô hình này phù hợp.
 Ta có thể chứng minh thông qua kiểm định F như sau:

Giải thuyết:

$H_0: R^2 = 0$ (Mô hình không phù hợp)

$H_1: R^2$ khác 0 (Mô hình phù hợp)

Kiểm định: Vì $\text{Prob}(F\text{-statistic}) = 0.000002 < \alpha = 5\%$. Nên bác bỏ H_0 .

Vậy mô hình phù hợp.

(4) Tìm thống kê F để thay vào dấu chấm hỏi của bảng kết xuất?

Áp dụng công thức đã học: $F_{tt} = \{R^2/(k-1)\}/\{(1-R^2)/(n-k)\}$

Với $k=4$, $n=29$, $R^2=0.684650$

Ta tính được: $F_{tt} = 18.09233$

(5) Viết công thức t bảng, biết $\alpha=5\%$, kiểm định 2 đuôi:

$$t^{\alpha/2}_{(n-k)} = t^{5\%/2}_{(29-4)}$$

(6) Không cần tra bảng, hãy viết công thức F bảng, biết $\alpha=5\%$?

$$\text{Công thức: } F^{\alpha}_{(k-1),(n-k)} = F^{5\%}_{(3),(25)}$$

(7) Giải thích ý nghĩa kinh tế của biến GNP trong mô hình?

Với điều kiện các yếu tố khác không đổi, khi GNP bình quân đầu người tăng lên 1 (đồng Lira) thì nhu cầu thuốc lá bình quân đầu người tăng 0.000804 (kg).

(8) Số liệu dùng để ước lượng hồi quy thuộc kiểu dữ liệu nào? Kiểu này thường vi phạm giả thiết gì khi xây dựng phương trình hồi quy tuyến tính (OLS)?

Dữ liệu thời gian.

Vi phạm giả thiết về $\text{Cov}(u_i, u_j)$ nên nguy cơ tự tương quan sẽ cao.

(9) Có người cho rằng với độ tin cậy 95%, ở giai đoạn 1960-1988, giá thuốc lá không ảnh hưởng đến nhu cầu tiêu thụ thuốc lá của người dân Thổ Nhĩ Kỳ. Quan điểm của bạn như thế nào (nếu sử dụng kết quả mô hình để giải thích)?

Theo kết quả của mô hình, với độ tin cậy 95%, ở giai đoạn 1960-1988, giá thuốc lá CÓ ảnh hưởng đến nhu cầu tiêu thụ thuốc lá của người dân Thổ Nhĩ Kỳ. Vì xác suất sai lầm của biến PRICE là rất thấp (0.0039), thấp hơn $\alpha=5\%$. Nên biến PRICE có ý nghĩa thống kê và có ảnh hưởng đến nhu cầu tiêu thụ thuốc lá (nếu trả lời bằng cách đặt giả thuyết kiểm định thống kê t cho hệ số hồi quy thì càng tốt).

(10) Cột Prob của bảng kết xuất có ý nghĩa:

Nói lên xác suất sai lầm khi giữ biến độc lập trong mô hình. Xác suất này càng lớn thì sai lầm càng cao, và ngược lại. Vì vậy, nếu $\text{Prob} < \alpha$ thì sai lầm khi giữ biến là dưới mức cho phép, hay nói cách khác, biến độc lập được giữ có ý nghĩa thống kê.

Bài 5 (Gồm 4 câu bên dưới)

Câu 1. Bảng kết xuất của một mô hình hồi quy tuyến tính như sau:

Dependent Variable: TIENGOI

Method: Least Squares

Date: 10/15/09 Time: 20:57

Sample: 1 50

Included observations: 50

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-44.32691	?	-0.702227	0.4859
THUNHAP	52.36692	11.13008	?	0.0000
R-squared	0.315624	Mean dependent var		210.7000
Adjusted R-squared	?	S.D. dependent var		273.6769
S.E. of regression	228.7509	Akaike info criterion		13.74232
Sum squared resid	2511696.	Schwarz criterion		13.81880
Log likelihood	-341.5581	F-statistic		?
Durbin-Watson stat	2.177023	Prob(F-statistic)		0.000022

Trong đó:

- TIENGOI : Tiền gửi tiết kiệm của khách hàng cá nhân (triệu đồng)
- THUNHAP: Thu nhập của khách hàng cá nhân, triệu đồng/tháng)

Yêu cầu: Tính toán số liệu để thay vào dấu chấm hỏi của bảng kết xuất trên?

Câu 2. (3 điểm) Giải thích ý nghĩa hệ số ước lượng được tô đậm trong mô hình sau đây:

$$\underline{\text{MH1.}} \quad Y = 234 - 27,14 * X_2 + \mathbf{11} * X_3 + u_i$$

$$\underline{\text{MH2.}} \quad \ln(Y) = 34,5 + 14,37 * X_2 + \mathbf{0,19} * X_3 + u_i$$

$$\underline{\text{MH3.}} \quad \ln(Y) = 12,45 - 0,52 * \ln(X_2) - \mathbf{0,11} * \ln(X_3) + u_i$$

$$\underline{\text{MH4.}} \quad Y = 450 + 129 * \ln(X_2) - \mathbf{117} * \ln(X_3) + u_i$$

Câu 3. (4 điểm) Hãy chọn một trong các chủ đề gợi ý dưới đây:

- a. Chi tiêu của hộ gia đình
- b. Điểm học tập của sinh viên
- c. Lợi nhuận của doanh nghiệp
- d. Chi phí chữa bệnh và ô nhiễm
- e. Khác (tự đề xuất)

Giả sử, bạn sẽ sử dụng phần mềm EVIEW và tiến hành xây dựng mô hình hồi quy có liên quan đến chủ đề vừa chọn. Vui lòng trả lời thật ngắn gọn các yêu cầu sau:

- (1) Biến phụ thuộc Y nào có thể sử dụng? (Nêu tên biến, đơn vị tính và diễn giải biến).
- (2) Những biến X_i bất kỳ nào có thể sử dụng như là các biến độc lập trong mô hình hồi quy để giải thích sự thay đổi của biến phụ thuộc đã phát biểu ở mục (1)? (Nêu rõ tên biến, diễn giải biến và đơn vị tính).
- (3) Dạng hàm toán học gì có thể ứng dụng cho việc xây dựng mô hình?
- (4) Nêu và giải thích kỳ vọng dấu của các hệ số ước lượng trong mô hình?

Giải bài 5

Câu 1. (3 điểm) Kết quả được điền vào như sau :

Dependent Variable: TIENGOI
Method: Least Squares
Date: 10/15/09 Time: 20:57
Sample: 1 50
Included observations: 50

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-44.32691	63.12336	-0.702227	0.4859
THUNHAP	52.36692	11.13008	4.704990	0.0000
R-squared	0.315624	Mean dependent var		210.7000
Adjusted R-squared	0.301367	S.D. dependent var		273.6769

S.E. of regression	228.7509	Akaike info criterion	13.74232
Sum squared resid	2511696.	Schwarz criterion	13.81880
Log likelihood	-341.5581	F-statistic	22.13693
Durbin-Watson stat	2.177023	Prob(F-statistic)	0.000022

Câu 2. (3 điểm) Giải thích ý nghĩa hệ số ước lượng được tô đậm trong mô hình sau đây:

- MH1.** $Y = 234 - 27,14 \cdot X_2 + \mathbf{11} \cdot X_3 + u_i$
 Với điều kiện các yếu tố khác không đổi, khi X_3 tăng lên 1 đơn vị thì Y sẽ tăng 11 đơn vị.
- MH2.** $\ln(Y) = 34,5 + 14,37 \cdot X_2 + \mathbf{0,19} \cdot X_3 + u_i$
 Với điều kiện các yếu tố khác không đổi, khi X_3 tăng lên 1 đơn vị thì Y sẽ tăng 19%.
- MH3.** $\ln(Y) = 12,45 - 0,52 \cdot \ln(X_2) - \mathbf{0,11} \cdot \ln(X_3) + u_i$
 Với điều kiện các yếu tố khác không đổi, khi X_3 tăng lên 1% thì Y sẽ giảm 0.11%.
- MH4.** $Y = 450 + 129 \cdot \ln(X_2) - \mathbf{117} \cdot \ln(X_3) + u_i$
 Với điều kiện các yếu tố khác không đổi, khi X_3 tăng lên 1% thì Y sẽ giảm 1.17 đơn vị.

Câu 3. (4 điểm)

Giả sử, bạn sẽ sử dụng phần mềm EVIEW và tiến hành xây dựng mô hình hồi quy có liên quan đến chủ đề Tiền gửi tiết kiệm của khách hàng cá nhân tại Ngân hàng Sacombank.

- (1) Biến phụ thuộc: Tiền gửi, ĐVT: triệu đồng.
- (2) Biến độc lập:
 - i. Thu nhập (triệu đồng/tháng). Kỳ vọng dấu +, vì thu nhập nhiều sẽ có khuynh hướng tiết kiệm nhiều.
 - ii. Độ tuổi (tuổi). Kỳ vọng dấu +, vì khách hàng cao tuổi có khuynh hướng tiết kiệm nhiều.
 - iii. Thủ tục (điểm). Được đo lường theo thang điểm do khách hàng chấm. Thấp nhất 1 điểm, cao nhất 5 điểm. Điểm cao thì thủ tục được khách hàng đánh giá tốt. Kỳ vọng dấu +, vì điểm cao có khuynh hướng thu hút khách hàng đến gửi và gửi nhiều tiền.
 - iv. Lãi suất (điểm). Được đo lường theo thang điểm do khách hàng chấm. Thấp nhất 1 điểm, cao nhất 5 điểm. Việc đánh giá dựa trên sự so sánh lãi suất của các ngân hàng có mặt trên địa bàn nghiên cứu. Điểm cao thì lãi suất cao, được khách hàng đánh giá tốt thì họ sẽ gửi nhiều.
- (3) Dạng hàm toán học tuyến tính: Linear

Bài 6. Kết xuất từ phần mềm EVIEW 7.0 như sau

a. Mô hình hồi quy gốc

Dependent Variable: Y
 Method: Least Squares
 Date: 11/26/10 Time: 09:06
 Sample: 1 50
 Included observations: 50

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	62.38597	1.269922	49.12584	0.0000
X2	0.149565	0.078642	1.901844	0.0635
X3	-0.000499	6.70E-05	-7.444031	0.0000
X4	0.000455	0.000235	1.938768	0.0587
R-squared	0.827575	Mean dependent var		64.33200
Adjusted R-squared	0.816329	S.D. dependent var		10.67537
S.E. of regression	4.575121	Akaike info criterion		5.955762

Sum squared resid	962.8595	Schwarz criterion	6.108723
Log likelihood	-144.8940	Hannan-Quinn criter.	6.014010
F-statistic	73.59401	Durbin-Watson stat	1.944382
Prob(F-statistic)	0.000000		

(1) Hãy đọc và nhận xét bảng kết xuất trên?

(2) Mô hình phù hợp không, các biến có ý nghĩa thống kê không, nếu độ tin cậy là 90%?

b. Mô hình hồi quy phụ

Heteroskedasticity Test: White

F-statistic	1.233529	Prob. F(9,40)	0.3027
Obs*R-squared	10.86241	Prob. Chi-Square(9)	0.2853
Scaled explained SS	8.964510	Prob. Chi-Square(9)	0.4406

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 11/26/10 Time: 09:06

Sample: 1 50

Included observations: 50

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	18.24191	16.24556	1.122886	0.2682
X2	-3.770434	1.852751	-2.035046	0.0485
X2^2	0.054301	0.047719	1.137936	0.2619
X2*x3	0.000131	0.000315	0.415996	0.6796
X2*x4	-6.71E-05	0.000174	-0.384453	0.7027
X3	-0.000119	0.001930	-0.061621	0.9512
X3^2	4.63E-09	3.58E-08	0.129363	0.8977
X3*x4	-5.33E-07	1.15E-06	-0.463625	0.6454
X4	0.013227	0.006940	1.905839	0.0639
X4^2	-4.27E-07	3.06E-07	-1.392955	0.1713

R-squared	0.217248	Mean dependent var	19.25719
Adjusted R-squared	0.041129	S.D. dependent var	27.16485
S.E. of regression	26.60036	Akaike info criterion	9.576583
Sum squared resid	28303.16	Schwarz criterion	9.958987
Log likelihood	-229.4146	Hannan-Quinn criter.	9.722204
F-statistic	1.233529	Durbin-Watson stat	1.753749
Prob(F-statistic)	0.302710		

Giải thích bảng trên? Kết luận của Anh/chị là gì?

Bài 7.

Mối quan hệ giữa tiền lương (TL), số năm kinh nghiệm làm việc (KN) và trình độ (DUM=1: Đại học; DUM=0: Trung học nghề) được thể hiện qua phương trình ước lượng bằng phương pháp OLS như sau:

$$TL = 1150 + 337*KN + 652*DUM$$

Giả định, mô hình này không bị vi phạm giả thiết, các biến độc lập đều có ý nghĩa thống kê, và mô hình này là phù hợp. *Yêu cầu:* Viết 2 phương trình ước lượng khi người lao động có trình độ đại học và có trình độ trung học nghề. Giải thích sự khác biệt về tiền lương theo trình độ của người lao động có số năm kinh nghiệm làm việc là 3 năm.

Tháng 09.2012

GV môn học: Trần Đức Luân

Email: tdluan@gmail.com

Tel: 0908.352.490