

Chương Trình Giảng Dạy Kinh tế Fulbright

Học kỳ Thu năm 2010

Các Phương Pháp Phân Tích Định Lượng

Gợi ý trả lời - Bài tập 7

HỒI QUY TUYẾN TÍNH ĐƠN II

Bài 1

Từ số liệu về tiêu thụ cà phê ở Mỹ (Y) và giá bán lẻ thực tế trung bình (X) ở Bảng 3.4

a. Ước lượng OLS

Bạn có thể tính toán thủ công từ các công thức; hoặc sử dụng công cụ **Tools\Data Analysis\Regression** trong Excel, có thể tính được bảng kết quả sau

<i>Regression Statistics</i>	
Multiple R	0.8141
R Square	0.6628
Adjusted R Square	0.6253
Standard Error	0.1287
Observations	11.0000

ANOVA					<i>Significance</i>	
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>F</i>	
Regression	1	0.2930	0.2930	17.6870	0.0023	
Residual	9	0.1491	0.0166			
Total	10	0.4421				

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	2.6911	0.1216	22.1269	0.0000	2.4160	2.9663
X	-0.4795	0.1140	-4.2056	0.0023	-0.7375	-0.2216

Từ bảng này, phương trình hồi quy như sau

$$\text{PRF: } Y_i = \beta_1 + \beta_2 X_i + u_i$$

$$\text{SRF: } Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}_i$$

Ước lượng điểm của β_1, β_2 theo phương pháp OLS tương ứng là $\hat{\beta}_1$ và $\hat{\beta}_2$. Với

$$\hat{\beta}_1 = 2.6911, \text{ và } \hat{\beta}_2 = -0.4795$$

b. Ý nghĩa của hệ số $\hat{\beta}_2$

$\hat{\beta}_2$ là tác động biên của lượng tiêu thụ cà phê trung bình ngày (Y) theo giá bán lẻ (X). Trong điều kiện các yếu tố khác không đổi, khi giá bán lẻ của một pound cà phê tăng 1 USD thì trung bình lượng cà phê tiêu thụ một ngày sẽ giảm 0.4795 tách, và ngược lại.

c. Sử dụng phương pháp khoảng tin cậy 95% để kiểm định giả thuyết cho rằng giá cà phê không có tác động gì đến mức tiêu dùng cà phê.

Chúng ta cần kiểm định giả thuyết $H_0: \beta_2=0$; và $H_1: \beta_2 \neq 0$

Khoảng tin cậy 95% của β_2 là (-0.7375, -0.2216). Số 0 nằm ngoài khoảng này nên bác bỏ H_0 , chấp nhận H_1 . Như vậy, giả thuyết “giá cà phê không có tác động gì đến mức tiêu dùng cà phê” đã bị bác bỏ (xét ở độ tin cậy 95%)

d. Kiểm định giả thuyết H_0 bằng cách sử dụng thống kê t

T-stat(X) = -4.2056 , $t_{0,025,9} = 2.2622$.

Do $|T\text{-stat}(X)| > t_{0,025,9}$ nên bác bỏ giả thuyết H_0

e. Kiểm định giả thuyết H_0 bằng cách sử dụng P-value

Xác suất mắc sai lầm loại I khi bác bỏ giả thuyết $H_0: \beta_2=0$ là 0.0023. P-value=0.0023 (< 0.05) nên ta bác bỏ giả thuyết H_0

f. Kiểm định giả thuyết H_0 bằng cách sử dụng thống kê F

Bảng ANOVA như sau

ANOVA					
	df	SS	MS	F	Significance F
Regression	k-1=1	ESS=0.2930	MSE=0.2930	F=17.6870	P-value(F)=0.0023
Residual	n-k=9	RSS=0.1491	MSR=0.0166		
Total	n-1=10	TSS=0.4421			

Giả thuyết H_0 ở trên cũng bị bác bỏ vì $F=17.6870 > F_{0,05,1,9}$ (với $F_{0,05,1,9}=5.1173$) hoặc P-value(F)= 0.0023 (<0.05)

g. Nếu giá bán lẻ trung bình một lb cà phê năm 1982 là 1.30 USD ($X_0=1.30$ USD), hãy ước lượng khoảng tin cậy 95% trung bình mỗi người dân Mỹ sẽ uống bao nhiêu tách cà phê mỗi ngày?

$$\hat{Y}_o = \hat{\beta}_1 + \hat{\beta}_2 X_o = 2.6911 - 0.4795 \times 1.30 = 2.0677$$

$$\sum x_i^2 = \frac{\hat{\sigma}^2}{\text{var}(\hat{\beta}_2)} = \frac{0.1287^2}{0.1140^2} = 1.2741$$

$$\text{var}(\hat{Y}_o) = \hat{\sigma}^2 \left[\frac{1}{n} + \frac{(X_o - \bar{X})^2}{\sum x_i^2} \right]$$

$$\text{var}(\hat{Y}_o) = 0.1287^2 \left[\frac{1}{11} + \frac{(1.3 - 1.0109)^2}{1.2714} \right] = 0.00259$$

$$SE(\hat{Y}_o) = \sqrt{\text{var}(\hat{Y}_o)} = 0.0509$$

Như vậy, khoảng tin cậy 95% của $E(Y|X_o)$ là $2.0677 \pm 2.2622 \times 0.0509$,
 hay (1.9525, 2.1829)

Bài 2

Xem các kết quả hồi qui sau đối với nền kinh tế Hoa Kỳ trong giai đoạn 1968-1987
 (\hat{Y} = chi tiêu của Hoa Kỳ đối với hàng hóa nhập khẩu và X = thu nhập khả dụng cá nhân, cả hai đều tính bằng tỷ USD, giá cố định 1982):

$$\hat{Y}_t = -261.09 + 0.2453X_t$$

Se (31.327) $(se(\hat{\beta}_2) = ?)$ $R^2 = 0.9388$
 t ($t_{\hat{\beta}_1} = ?$) (16.616) n = 20

Câu a: Điền các số vào ô trống trong ngoặc?

Trong kết quả trên, $t_{\hat{\beta}_1} = \frac{\hat{\beta}_1}{se_{\hat{\beta}_1}} = \frac{-261.09}{31.327} = -8.3343$

và $se(\hat{\beta}_2) = \hat{\beta}_2 / t_{\hat{\beta}_2} = 0.2453 / 16.616 = 0.0148$

Câu b: Bạn giải thích hệ số 0.2453 như thế nào? Và hệ số -261.09?

$\hat{\beta}_2 = 0.2453 > 0$ nên chi tiêu của Hoa Kỳ đối với hàng nhập khẩu đồng biến với thu nhập khả dụng cá nhân. Con số này cũng cho biết: với điều kiện các yếu tố khác không đổi, khi thu nhập khả dụng cá nhân tăng (giảm) 1 tỷ USD thì chi tiêu đối với hàng nhập khẩu tăng (giảm) trung bình 0.2453 tỷ USD.

Về mặt toán học, $\hat{\beta}_1 = -261.09$ cho biết khi thu nhập khả dụng cá nhân = 0, chi tiêu cho hàng nhập khẩu của Hoa Kỳ là -261.09 tỷ USD. Tuy nhiên về mặt kinh tế, người ta ít quan tâm đến ý nghĩa kinh tế của hệ số chặn này.

Câu c: Kiểm định giả thuyết cho rằng hệ số góc bằng 0 với mức ý nghĩa 5%? p-value của trị thống kê kiểm định bằng bao nhiêu?

Đặt giả thuyết $H_0: \beta_2 = 0$

Và giả thuyết thay thế $H_1: \beta_2 \neq 0$

$$t_{\hat{\beta}_2} = 16.616$$

Với mức ý nghĩa 5% ta có, $t_{0.025,18} = 2.1009$

Như vậy $|t_{\hat{\beta}_2}| > t_{0.025,18}$ nên bác bỏ H_0 , chấp nhận H_1 . Hay có thể khẳng định hệ số góc

β_2 khác 0 ở mức ý nghĩa 5%.

Giá trị p_{value} của trị thống kê kiểm định được tính bằng hàm Tdist trong Excel.

$P_{\text{value}} = \text{Tdist}(16.616, 18, 2) = 2.2998 \cdot 10^{-12}$ (con số này rất nhỏ so với 0.05)