

Chương Trình Giảng Dạy Kinh tế Fulbright

Học kỳ Thu năm 2010

Các Phương Pháp Phân Tích Định Lượng

Đáp án Bài tập 6

HỘI QUY ĐƠN BIẾN

Bài 2.11

a. $Y_i = \beta_1 + \beta_2 \left(\frac{1}{X_i} \right) + u_i$

b. $Y_i = \beta_1 + \beta_2 \ln X_i + u_i$

c. $\ln Y_i = \beta_1 + \beta_2 X_i + u_i$

d. $\ln Y_i = \ln \beta_1 + \beta_2 \ln X_i + u_i$

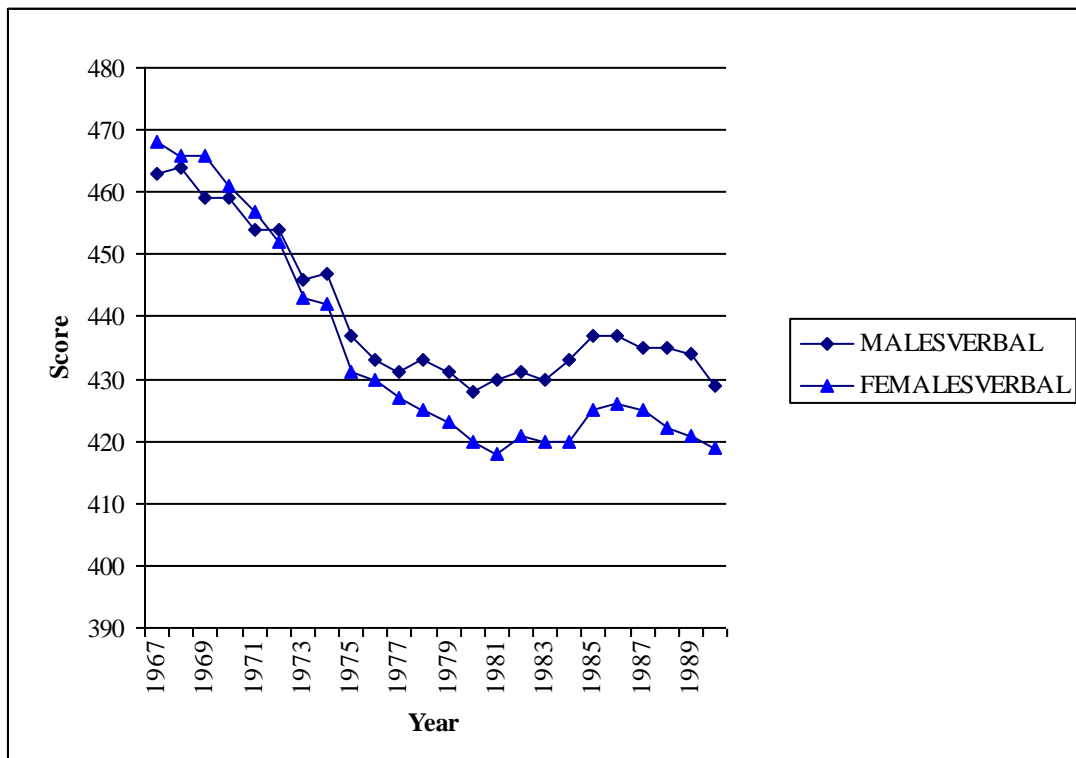
e. $\ln Y_i = \beta_1 - \beta_2 \left(\frac{1}{X_i} \right) + u_i$

Mô hình a, b, c, và e: là các mô hình tuyến tính theo tham số, phi tuyến theo biến số. Chúng đều là mô hình hồi quy tuyến tính.

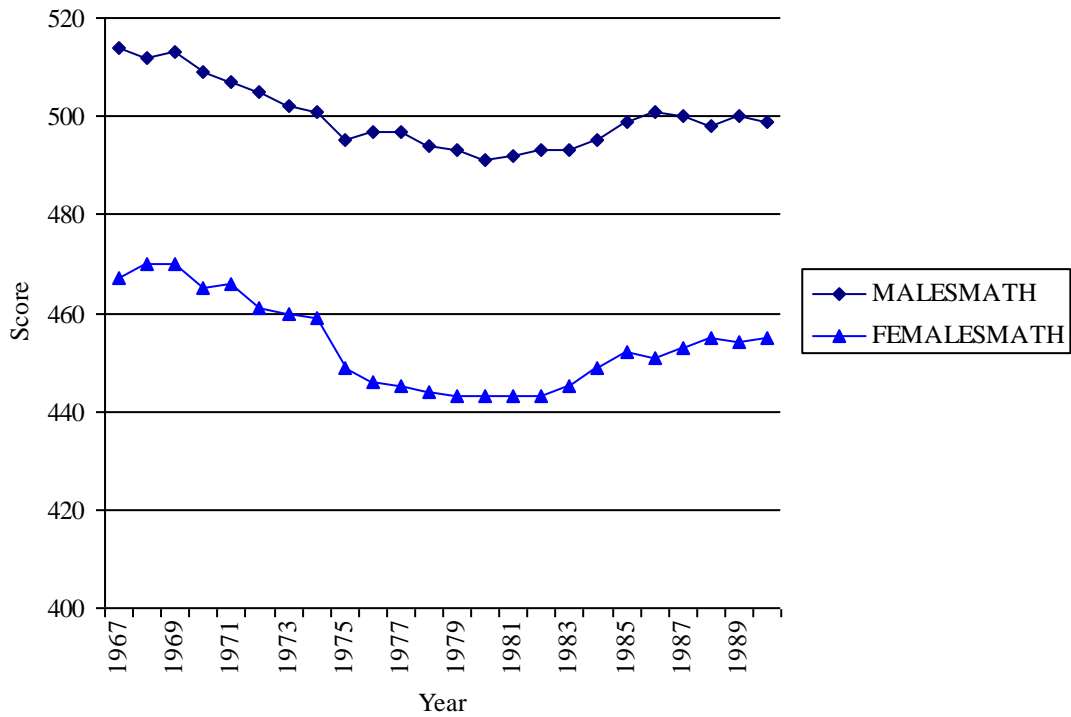
Mô hình d. phi tuyến theo biến số; nếu đặt $\alpha = \ln(\beta_1)$ thì mô hình này sẽ là mô hình tuyến tính theo tham số, và là mô hình hồi quy tuyến tính.

Bài 2.16

a. Đồ thị điểm vắn đáp của nam và nữ



Đồ thị điểm toán của nam và nữ

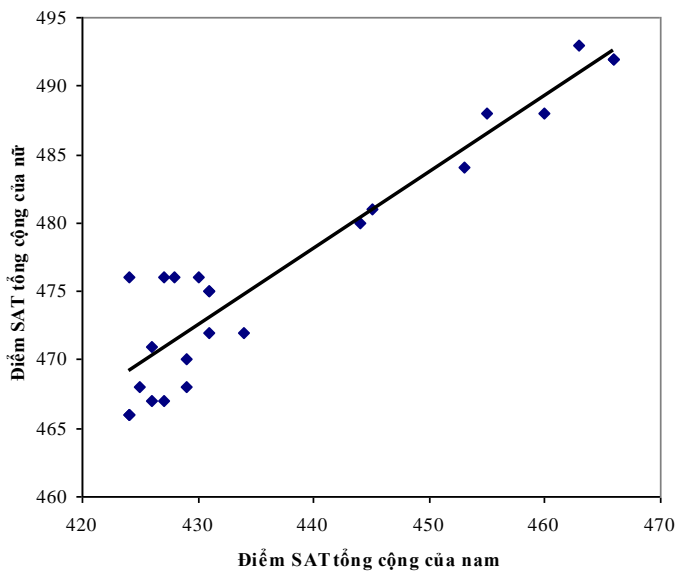


b. Nhận xét

Điểm toán của nam và nữ có khuynh hướng giảm xuống qua các năm, giảm ở mức thấp nhất vào năm 1980. Khuynh hướng tương tự cũng xảy ra đối với điểm văn đáp của cả nam và nữ.

c. Khi biết điểm văn đáp của nam và nữ, để dự đoán điểm toán của họ, chúng ta có thể xây dựng hàm hồi quy điểm toán phụ thuộc vào điểm văn đáp cho nam, và cho nữ.

d. Đồ thị phân tán



Đồ thị trên cho thấy, điểm SAT tổng cộng của nam có mối tương quan đồng biến với điểm SAT tổng cộng của nữ. Mối quan hệ giữa hai biến này theo dạng đường thẳng

Bài 3.9

a.

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

$$\hat{\alpha}_1 = \bar{Y} - \hat{\beta}_2 \bar{x} = \bar{Y} \text{ (vì } \sum x_i = 0)$$

$$\text{var}(\hat{\beta}_1) = \frac{\sum X_i^2}{n \sum x_i^2} \sigma^2$$

$$\text{var}(\hat{\alpha}_1) = \frac{\sum x_i^2}{n \sum x_i^2} \sigma^2 = \frac{\sigma^2}{n}$$

Vì vậy, ước lượng hệ số chặn và phương sai cho hệ số chặn ở mô hình 1 và mô hình 2 không bằng nhau

b.

$$\hat{\beta}_2 = \frac{\sum y_i x_i}{\sum x_i^2}$$

$$\hat{\alpha}_2 = \frac{\sum [(X_i - \bar{X}) - (\overline{X_i - \bar{X}})] (Y_i - \bar{Y})}{\sum [(X_i - \bar{X}) - (\overline{X_i - \bar{X}})]^2} = \frac{\sum y_i x_i}{\sum x_i^2}$$

Dễ dàng xác định được

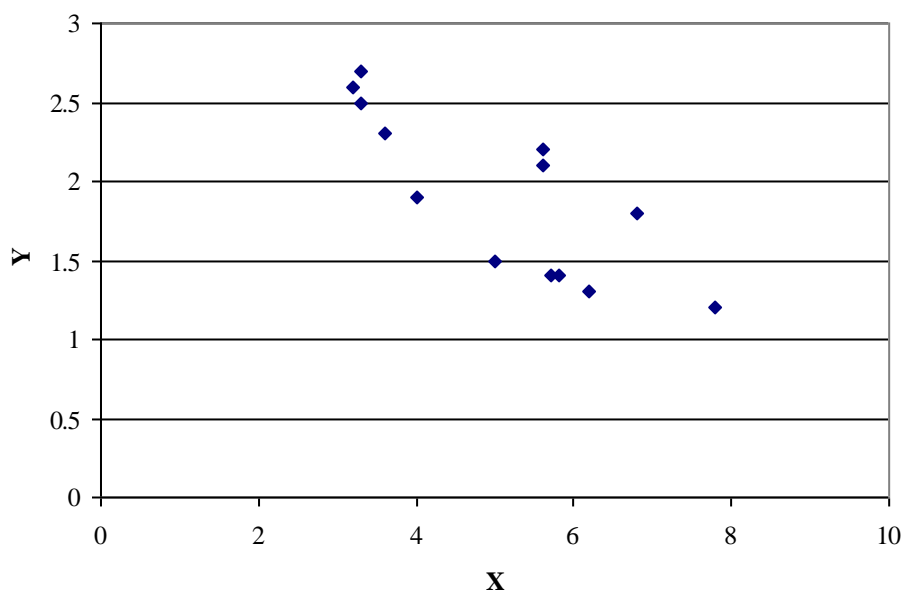
$$\text{var}(\hat{\beta}_2) = \text{var}(\hat{\alpha}_2) = \frac{\sigma^2}{\sum [(X_i - \bar{X}) - (\overline{X_i - \bar{X}})]^2} = \frac{\sigma^2}{\sum (x_i - \bar{x})^2} = \frac{\sigma^2}{\sum x_i^2}$$

Do vậy, ước lượng hệ số độ dốc, và phương sai của hệ số độ dốc ở mô hình 1, và mô hình 2 là như nhau

c. Mô hình 2 có thể dễ tính toán hơn mô hình 1, tuy nhiên hiện nay với sự hỗ trợ của máy tính thì lợi thế này không đáng quan tâm.

Bài 3.17

a.



b.

Sử dụng công cụ Tools\Data Analysis\Regression của Excel bạn có thể tìm được kết quả sau

SUMMARY OUTPUT

<i>Regression Statistics</i>	
Multiple R	0.808
R Square	0.653
Adjusted R Square	0.622
Standard Error	0.322
Observations	13

ANOVA

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	2.153	2.153413	20.71477	0.0008
Residual	11	1.144	0.103955		
Total	12	3.297			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	3.366	0.331	10.17	0.000	2.638	4.095
X	-0.286	0.063	-4.55	0.001	-0.425	-0.148

Từ Kết quả này, ta có thể xác định được phương trình hồi quy mẫu

$$Y_i = b_1 + b_2 X_i + \hat{u}_i$$

$$Y_i = 3.366 - 0.286X_i + \hat{u}_i$$

$$SE(b_1) = 0.331, SE(b_2) = 0.063$$

Hệ số xác định $R^2 = 0.653$

Do X và Y có mối liên hệ nghịch chiều (hệ số $b_2 < 0$) nên hệ số tương quan giữa X và Y = $-\sqrt{R^2} = -0.808$

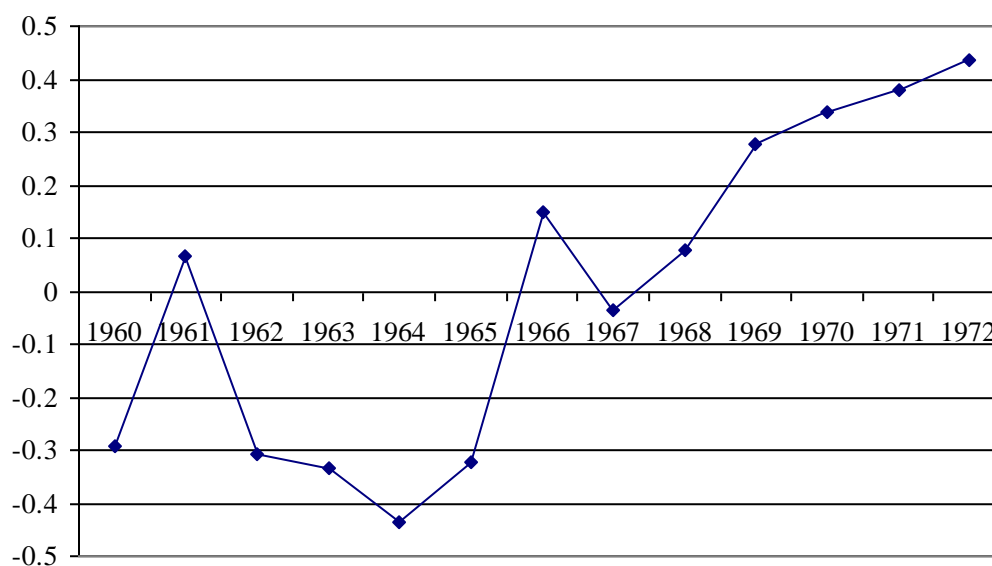
d.

Hàm hồi quy mẫu cho biết, X ảnh hưởng nghịch chiều lên Y vì b_2 mang dấu âm. Trong điều kiện các yếu tố khác không đổi, nếu X (tỷ lệ thất nghiệp) tăng 1 đơn vị thì trung bình Y (tỷ lệ bỏ việc trong 100 công nhân) giảm 0.286 đơn vị, và ngược lại.

Biến X giải thích được 65.3% biến thiên của biến Y

e. Vẽ đồ thị của phần dư

Residuals



Phần dư có những giá trị âm, có những giá trị dương biểu thị sự chênh lệch giữa giá trị thực tế và giá trị ước lượng. Trong mô hình này, phần dư tăng dần theo thời gian.

f.

Kết quả ước lượng phương trình hồi quy với dữ liệu từ 1966-1978 khác với kết quả ước lượng phương trình hồi quy ở câu b. Sự khác biệt này là do dữ liệu ở 2 mẫu khác nhau. Đường hồi quy ở câu f gần với các điểm dữ liệu trên đồ thị phân tán hơn đường hồi quy ở câu b do R^2 tốt hơn.