Chương Trình Giảng Dạy Kinh tế Fulbright

Học kỳ Thu năm 2012

Các Phương Pháp Phân Tích Định Lượng

LỜI GIẢI ĐỀ NGHỊ BÀI TẬP 6

HỒI QUY ĐƠN BIẾN

Ngày Phát: Thứ hai 12/11/2012

Ngày Nộp: Thứ hai 26/11/2012

Bản in nộp lúc **8h20 sáng**, tại Hộp nộp bài tập trong phòng Lab

Bản điện tử gửi lên http://intranet.fetp.edu.vn:81

Bài 1:

Trong các mô hình sau đây, mô hình nào là mô hình tuyến tính. Đối với các mô hình không tuyến tính, hãy thử dùng các phép biến đổi toán học thích hợp để chuyển chúng về dạng tuyến tính

a. $Y_i = \beta_1 + \beta_2 \dfrac{1}{X_i}$ là mô hình này tuyến tính theo tham số

b. $Y_i = \beta_1 + \beta_2 \ln(X_i)$ là mô hình này tuyến tính theo tham số

c. $Y_i = \beta_1 + e^{\beta_2 X_i}$ mô hình này không phải là mô hình tuyến tính do các hệ số bêta phi tuyến. Phép biến đổi như sau:

$Y_i = \beta_1 + e^{\beta_2 X_i} \Rightarrow Y_i - \beta_1 = e^{\beta_2 X_i}$ lấy ln 2 vế ta được $\ln(Y_i - \beta_1) = \ln(e^{\beta_2 X_i}) \Rightarrow \ln(Y_i - \beta_1) = \beta_2 X_i$,

đặt $\ln(Y_i - \beta_1)$ =Y_i' ta có hàm hồi quy tuyến tính Y_i'= $\beta_2 X_i$

d. $Y_i = e^{\beta_1 + \beta_2 X_i} \Rightarrow \ln(Y_i) = \ln(e^{\beta_1 + \beta_2 X_i}) \Rightarrow Y_i = (\beta_1 + \beta_2 X_i)\ln e = \beta_1 + \beta_2 X_i$

Như vậy sau khi biến đổi mô hình (d) trở thành mô hình tuyến tính có dạng Y_i= $\beta_1 + \beta_2 X_i$

e. $Y_i = \dfrac{1}{1 + e^{\beta_1 + \beta_2 X_i}}$. mô hình này không phải là mô hình tuyến tính do các hệ số bêta phi tuyến. Phép biến đổi như sau:

$Y_i = \dfrac{1}{1 + e^{\beta_1 + \beta_2 X_i}} \Rightarrow \dfrac{1}{Y_i} = 1 + e^{\beta_1 + \beta_2 X_i} \Rightarrow \dfrac{1}{Y_i} - 1 = e^{\beta_1 + \beta_2 X_i}$, lấy ln 2 vế ta được

$\ln(\dfrac{1}{Y_i} - 1) = \ln(e^{\beta_1 + \beta_2 X_i}) \Rightarrow \ln(\dfrac{1}{Y_i} - 1) = \beta_1 + \beta_2 X_i$

Đặt $\ln(\dfrac{1}{Y_i} - 1)$ = Y_i' ta có mô hình tuyến tính Y_i' = $\beta_1 + \beta_2 X_i$

f. $Y_i = \dfrac{X_i}{\beta_1 + \beta_2 X_i}$ mô hình này không phải là mô hình tuyến tính do các hệ số bêta phi tuyến. Phép biến đổi như sau:

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

Đặt $\frac{1}{Y_i}$ = Y_i' ta có mô hình tuyến tính Y_i' $= \beta_2 + \beta_1 \frac{1}{X_i}$

g. $Y_i = \beta_1 + (\beta_2)^2 X_i$ (Không tuyến tính)

Đặt $\beta_2^2 = \beta_2'$ ta có $Y_i = \beta_1 + (\beta_2)' X_i$

h. $Y_i = \beta_1 + (\beta_1) e^{\beta_2 X_i}$ mô hình này không phải là mô hình tuyến tính do các hệ số bêta phi tuyến. Phép biến đổi như sau:

$Y_i = \beta_1 + (\beta_1) e^{\beta_2 X_i} \Leftrightarrow \dfrac{Y_i - \beta_1}{\beta_1} = e^{\beta_2 X_i}$ lấy ln 2 vế ta được

$$\ln\left(\frac{Y_i - \beta_1}{\beta_1}\right) = \ln(e^{\beta_2 X_i}) \Rightarrow \ln\left(\frac{Y_i - \beta_1}{\beta_1}\right) = \beta_2 X_i$$

Đặt $\ln\left(\dfrac{Y_i - \beta_1}{\beta_1}\right)$ = Y_i' ta có mô hình tuyến tính Y_i'= $\beta_2 X_i$

i. $\ln Y_i = \beta_1 + \beta_2 X_i$ là mô hình tuyến tính thep các hệ số bêta

j. $\ln Y_i = \beta_1 + \ln(\beta_2 X_i)$ mô hình này không phải là mô hình tuyến tính do các hệ số bêta phi tuyến.
Phép biến đổi như sau:
=> $\ln Y_i = \beta_1 + \ln \beta_2 + \ln X_i$ đặt $ln\beta_2 = \alpha$ ta được $\ln Y_i = \beta_1 + \alpha + \ln X_i$

Bài 2. Chứng minh rằng: Một khi đã thu được các ước lượng bình phương tối thiểu thông thường (OLS) từ dữ liệu mẫu thì:
a.Đường hồi quy mẫu đi qua điểm có tọa độ là giá trị trung bình mẫu
Cách 1:

Ta cần chứng minh $\widehat{Y_i} = E(\hat{Y}_i / X_i) - \hat{u}_i$

$\hat{\beta}_2 = [\sum(X_i - \bar{X}).\sum(Y_i - \bar{Y})] / \sum(X_i - \overline{X})^2$
$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}$

$\widehat{Y_i} = \widehat{\beta_1} + \widehat{\beta_2}$ Xi + $\hat{u}_i$ => $\sum Y_i = \sum \widehat{\beta_1} + \sum \widehat{\beta_2}$ .Xi +$\sum \hat{u}_i$

=>$\sum \hat{u}_i = \sum Y_i - \sum \widehat{\beta_1} - \sum \widehat{\beta_2}$ .Xi, vì $\hat{u}_i = 0$ nên= $\sum Y_i - n.\widehat{\beta_1} - \sum \widehat{\beta_2}$ .Xi=0 với $n.\widehat{\beta_1} = \sum \widehat{\beta_1}$.

⇨ $\frac{1}{n}\sum Y_i = \widehat{\beta_1} + \widehat{\beta_2}\frac{1}{n}\sum$ Xi

⇨ $\bar{Y} = \widehat{\beta_1} + \widehat{\beta_2}\bar{X}$ (với $\bar{Y} = \frac{1}{n}\sum Y_i$ và $\bar{X} = \frac{1}{n}\sum$ Xi)

Như vậy đường hồi quy mẫu đi qua giá trị trung bình

Cách 2:

Ta có:

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

⇔ $\frac{\sum Y_i}{n} = \frac{\sum \hat{\beta}_1 + \hat{\beta}_2 \sum X_i + \sum \hat{U}_i}{n}$

⇔ $\bar{Y} = \hat{\beta}_1 + \hat{\beta}_2 X_i + (\sum \hat{U}_i)/n$

Mà $(\sum \hat{U}_i)/n = 0$

⇨ $\bar{Y} = \widehat{\beta_1} + \widehat{\beta}_2 \bar{X}$

Hay đường hồi quy mẫu đi qua điểm có tọa độ là giá trị trung bình mẫu

b.Giá trị trung bình của ước lượng bằng với giá trị trung bình của mẫu $\bar{\hat{Y_\iota}} = \bar{Y}$

Ta có: $\sum \hat{u_\iota} = 0$

$\sum(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i) = 0$

$\sum(Y_i - \hat{Y}_i) = 0$

$\sum Y_i = \sum \hat{Y}_i \rightarrow \bar{\hat{Y_\iota}} = \bar{Y}_i$ (đpcm)

c.Trung bình phần dư bằng 0, E $(\hat{u}_i) = 0$

$\sum \hat{u}_i = \sum (Y_i - \bar{Y_\iota})^2 = f(\beta_1, \beta_2) = \sum (Yi - \widehat{\beta_1} - \widehat{\beta_2}\, Xi)^2$

Phương pháp OLS ta có $\widehat{\beta_1}$ và $\widehat{\beta_2}$ sao cho để $\sum \hat{u}_i{}^2$ nhỏ nhất.

Hay: $\frac{\partial f}{\partial \widehat{\beta} 2} = -2\sum (Yi - \widehat{\beta_1} - \widehat{\beta_2}\, Xi) = 0 \Rightarrow \sum [Yi - (\widehat{\beta_1} + \widehat{\beta_2}\, Xi)] = 0$

$=> \sum \hat{u}_i = 0$ nên E $(\hat{u}_i) = 0$

d.Phần dư $\hat{u}_i$ không tương quan với giá trị dự báo Y_i

Ta có: $\mathrm{cov}(\hat{u}_i, Y_i) = E((\hat{u}_i - \bar{\hat{u}}_i)(Y_i - \bar{Y}))$

Với: $\bar{\hat{u}}_i = \frac{\sum \hat{u}_i}{n} = \frac{0}{n} = 0$

Nên $\mathrm{cov}(\hat{u}_i, Y_i) = E[(\hat{u}_i(Y_i - \bar{Y})] = E(\hat{u}_i Y_i - \hat{u}_i \bar{Y})$

$= E(\hat{u}_i Y_i) - \bar{Y} E(\hat{u}_i) = E(\hat{u}_i Y_i)$

$= E[\hat{u}_i(\hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}_i)]$

$= \hat{\beta}_1 E(\hat{u}_i) + \hat{\beta}_2 E(\hat{u}_i X_i) + E^2(\hat{u}_i) = 0$

e.Phần dư $\hat{u}_i$ không tương quan với giá trị X_i

Ta có: cov($\hat{u}_i, X_i$) = E[$(\hat{u}_i - \bar{\hat{u}}_i)(X_i - \bar{X})$]

Ta có: $\bar{\hat{u}}_i = \frac{\sum \hat{u}_i}{n}$ = $\frac{0}{n}$ = 0

Nên cov($\hat{u}_i, X_i$) = E[$(\hat{u}_i(X_i - \bar{X})$]

= E $(\hat{u}_i X_i) - \bar{X}E(\hat{u}_i)$

= 0 – 0 = 0

Bài 3.

a. Bạn hãy cho biết mối quan hệ giữa chi tiêu và thu nhập.

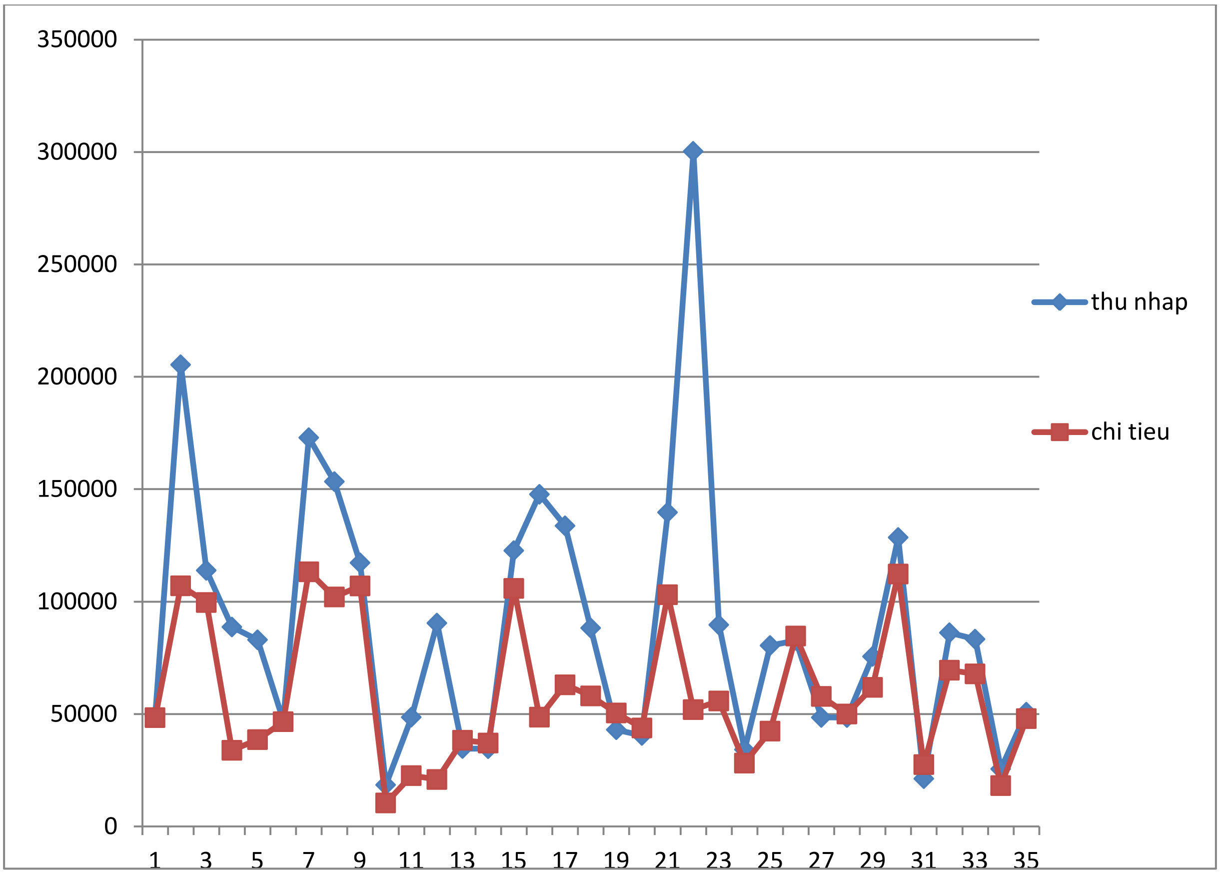
Thu nhập và chi tiêu có mối quan hệ đồng biến với nhau. Nếu thu nhập tăng thì chi tiêu tăng, thu nhập giảm thì chi tiêu giảm.

Lý thuyết về nền kinh tế thực chứng minh cho điều này. Lý thuyết này cho thấy mối tương quan giữa thu nhập và chi tiêu.

Hàm tiêu dùng C = C_o + Cm.Yd, trong đó Yd =(Y-T) (thu nhập khả dụng bằng thu nhập trừ đi thuế), Cm hay MPC là khuynh hướng tiêu dùng biên phản ánh mức thay đổi của tiêu dùng khi thu nhập khả dụng thay đổi 1 đơn vị. MPC = $\frac{\Delta C}{\Delta Yd}$

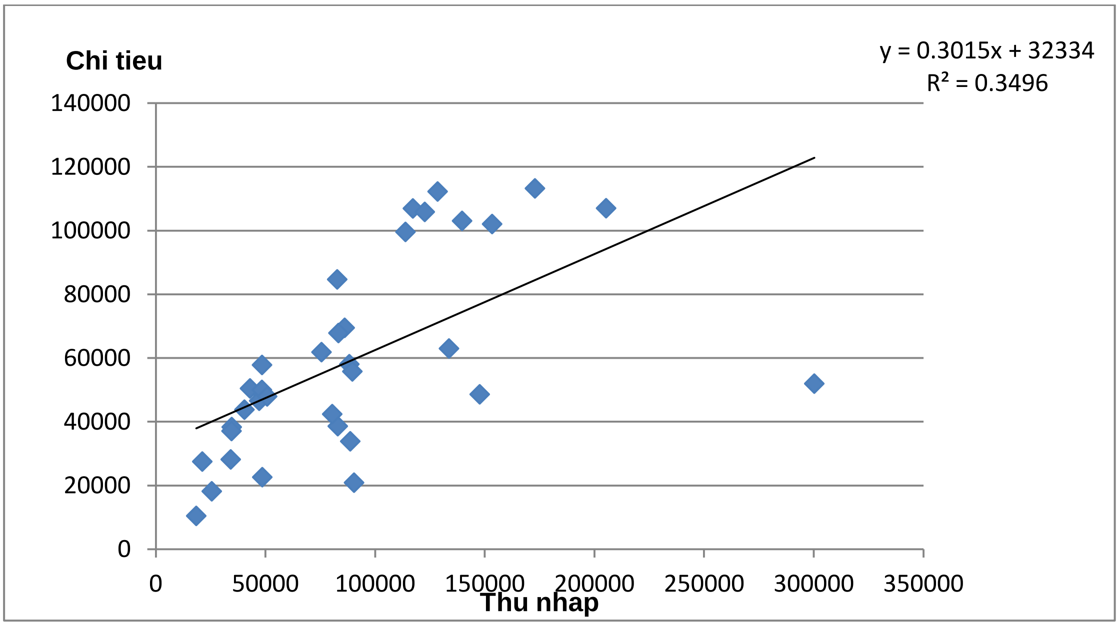
b. Khai thác bộ dữ liệu VHLSS 2008, hãy chọn ra 35 quan sát tính toán các số liệu về thu nhập cũng như chi tiêu của các quan sát này, thể hiện sự biến động của hai tập dữ liệu này trên cùng một đồ thị.

Tùy vào mẫu ngẫu nhiên mà người làm khai thác trong VHLSS 2008 sẽ cho đồ thị thể hiện sự biến động của tập dữ liệu. Tuy nhiên đồ thị phải thể hiện được mối quan hệ đồng biến trên cùng một đồ thị.



c.Giả sử chúng ta sử dụng hàm hồi quy tuyến tính, viết phương trình mô tả mối quan hệ giữa chi tiêu và thu nhập

Theo dữ liệu khai thác, ta vẽ đồ thị và viết phương trình mô tả mối quan hệ giữa chi tiêu và thu nhập như sau:



Hàm hồi quy tuyến tính giữa thu nhập và chi tiêu là

$\widehat{Y_i} = \widehat{\beta_1} + \widehat{\beta_2}$ X_i = 32.334+ 0,3015* X_i

Kỳ vọng về dấu của biến: theo lý thuyết thì việc chi tiêu có mối quan hệ đồng biến với thu nhập nên ta kỳ vọng $\widehat{\beta_2} > 0$

d. Ước lượng hệ số tiêu dùng biên MPC

Tùy theo dữ liệu để ước lượng hệ số MPC hay chính là ước lượng hệ số $\widehat{\beta_2}$

Với $\widehat{\beta_2}$ thỏa mãn điều kiện $0 <\widehat{\beta_2}<1$.

Theo cơ sở dữ liệu ngẫu nhiên thì ta có:

MPC = $\hat{\beta}_2 = [\sum(X_i - \bar{X}).\sum(Y_i - \bar{Y})]/\sum(X_i - \overline{X})^2 = 0{,}3015$

e. Giả định để câu d có độ tin cậy cao

- Mô hình là hồi quy tuyến tính.
- Các giá trị của thu nhập được cố định trong việc lấy mẫu lặp lại.
- Kỳ vọng của phần dư bằng 0.
- Đồng phương sai giữa các giá trị phần dư và thu nhập tương ứng bằng 0.
- Có sự biến thiên trong các giá trị của thu nhập.
- Phương sai của sai số không đổi.
- Có sự độc lập theo chuỗi. Không có tương quan giữa các sai số.
- Mô hình hồi quy được xác định một cách đúng đắn.
- Không có tính đa cộng tuyến hoàn toàn.
- Số mẫu lớn hơn số biến độc lập

Bài 4. Tập đoàn du lịch Woody có một hệ thống gồm rất nhiều nhà hàng ở rất nhiều địa phương. Tập đoàn này đã thu thập dữ liệu mẫu gồm 33 nhà hàng. Thông tin của mỗi nhà hàng được lưu trữ trong các biểu sau:

- Y: Số lượt khách hàng được phục vụ trong 1 năm (lượt khách)
- I: Thu nhập trung bình một hộ của những hộ ở quanh nhà hàng trong phạm vi bán kính 3 dặm (USD)
- P: Số người dân sống xung quanh nhà hàng trong phạm vi bán kính 3 dặm (người)
- N: Số đối thủ cạnh tranh trực tiếp với nhà hàng trong phạm vi bán kính 2 dặm.

a.Dựa vào khung lý thuyết cung cầu để thiết lập mô hình về các yếu tố ảnh hưởng đến Y.

Lý thuyết cung cầu cho thấy cầu của một hàng hóa phụ thuộc vào thu nhập của người tiêu dùng, thị hiếu của người tiêu dùng, giá kỳ vọng, số người mua, giá hàng thay thế, giá hàng bổ sung.

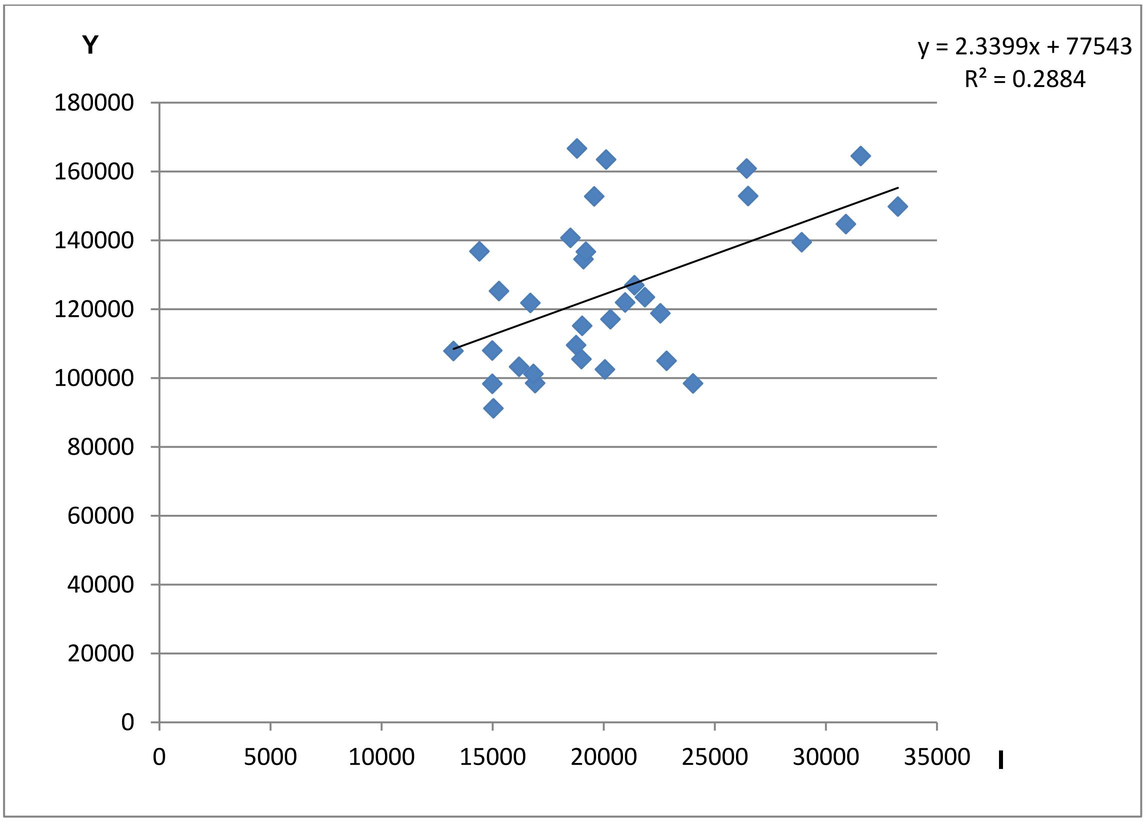
- Kỳ vọng về mối quan hệ giữa I và Y là: nếu thu nhập trung bình một hộ của những hộ xung quanh nhà hàng trong phạm vi bán kính 3 dặm tăng thì số lượng khách hàng được phục vụ trong một năm sẽ tăng. Tức là mối quan hệ giữa I và Y là đồng biến.

- Kỳ vọng về mối quan hệ giữa P và Y là: nếu số người dân sống xung quanh nhà hàng trong phạm vi bán kính 3 dặm (người) nhiều thì khả năng số lượt khách hàng được phục vụ trong một năm sẽ nhiều. Tức là mối quan hệ giữa P và Y là đồng biến.

- Kỳ vọng về mối quan hệ giữa N và Y là: nếu số đối thủ cạnh tranh trực tiếp với nhà hàng trong phạm vi bán kính 2 dặm tăng thì khả năng số lượt khách hàng được phục vụ trong một năm sẽ giảm. Tức là mối quan hệ giữa N và Y là nghịch biến.

b.Vẽ biểu đồ phân tán, tính hệ số tương quan giữa Y và I, Y và P, giữa Y và N và cho biết nhận định của bạn về mối quan hệ giữa các biến này.

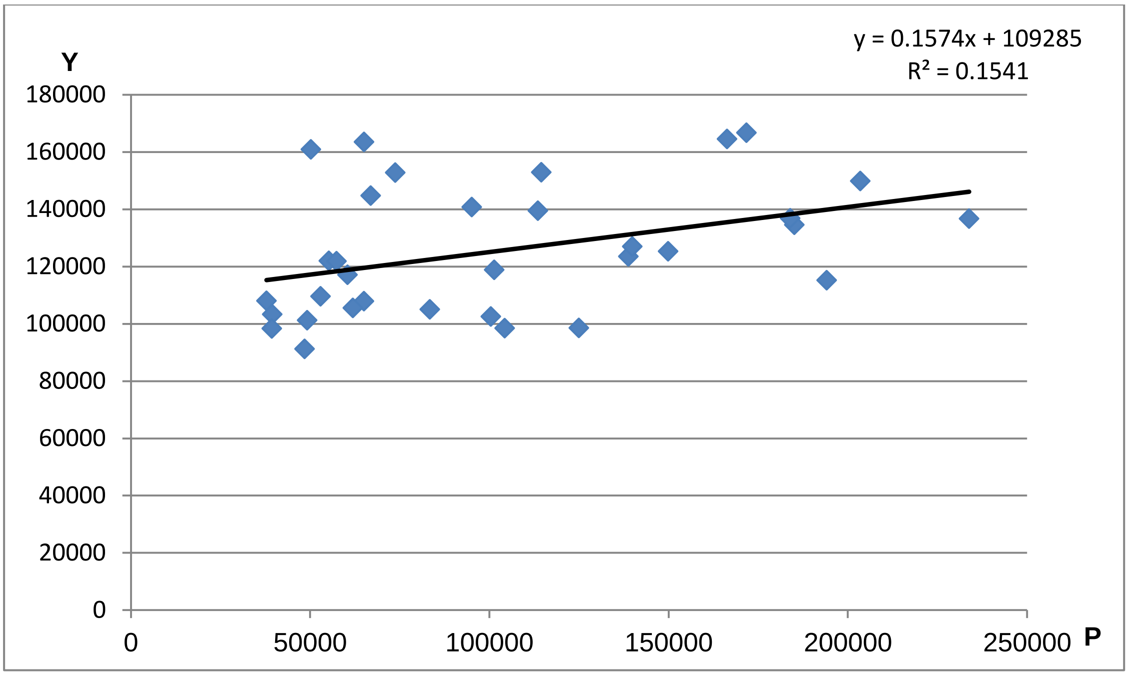
Biểu đồ phân tán biểu thị mối quan hệ giữa Y và I



Sử dụng công cụ trong Excel ta vẽ biểu đồ phân tán và tính được hệ số tương quan của Y và I $r_{Y/I} = \sqrt{R^2_{Y/I}} = \sqrt{0{,}2884} \approx 0{,}5367$

Nhận định về mối quan hệ Y và I: Số lượt khách hàng được phục vụ trong 1 năm (Y) và thu nhập trung bình một hộ của những hộ ở quanh nhà hàng trong phạm vi bán kính 3 dặm (I) có mối quan hệ đồng biến và tương đối chặt với nhau (R = 0,5367)

Biểu đồ phân tán biểu thị mối quan hệ giữa Y và P

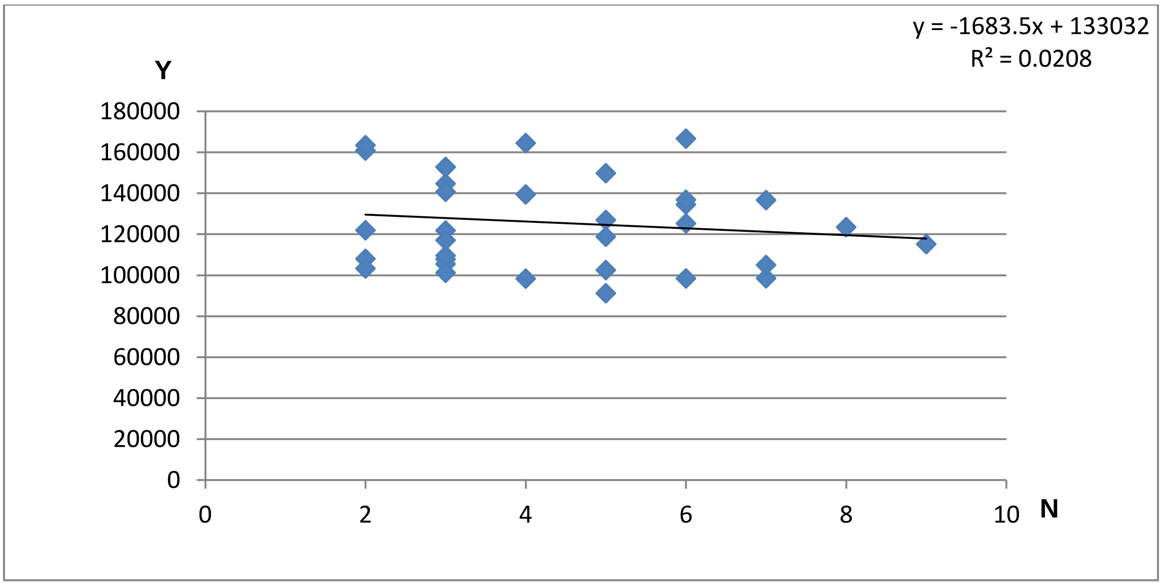


Sử dụng công cụ trong Excel ta vẽ biểu đồ phân tán và tính được hệ số tương quan của Y và P

$r_{Y/P} = \sqrt{R^2_{Y/P}} = \sqrt{0{,}154} \approx 0{,}3924$

Nhận định về mối quan hệ Y và P: Số lượt khách hàng được phục vụ trong 1 năm (Y) và số người dân sống xung quanh nhà hàng trong phạm vi bán kính 3 dặm (P) có mối quan hệ đồng biến và tương đối không chặt với nhau (R = 0,3924)

Biểu đồ phân tán biểu thị mối quan hệ giữa Y và N



Sử dụng công cụ trong Excel ta vẽ biểu đồ phân tán và tính được hệ số tương quan của Y và N

$R_{Y/N} = -\sqrt{R^2_{Y/N}} = -\sqrt{0{,}020} \approx -0{,}141$

Như vậy giữa Y và N có quan hệ nghịch biến và tương quan yếu vì r = -0,167. Số lượt khách hàng được phục vụ trong năm chịu ảnh hưởng của số đối thủ cạnh tranh trực tiếp.

c.Hãy sử dụng các công thức để tính toán các hệ số hồi quy trong hàm hồi quy tuyến tính Y = f(P)

Tính các số liệu cần thiết: $\bar{P} = \frac{\sum P_i}{n} = \frac{3.428.287}{33} \approx 103.887$ (người)

$\bar{Y} = \frac{\sum Y_i}{n} = \frac{4.145.942}{33} \approx 125.635$ (lượt khách được phục vụ)

Sử dụng các công thức và các số liệu tính toán ở bảng trên để xác định các ước lượng tham số của phương trình như sau:

$\hat{\beta}_2 = \frac{\sum p_i y_i}{\sum y_i^2} = \frac{15.728.356.029}{99.938.509.616} \approx 0,1574$

$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2\bar{P} = 125.635 – 0,1574 \times 103.887 = 109.284,759$

Vậy có thể ước lượng hàm hồi quy tuyến tính Y=f(P) như sau: $Y_i = 109.284,759 + 0,1574P_i + \hat{u}_i$

d.Hãy sử dụng các công thức để tính toán phương sai, sai số chuẩn của các hệ số hồi quy trong hàm hồi quy tuyến tính Y = f(P)

Nhà hàng	**Y**	**P**	**$y_i = Y_i - \bar{Y}$ (Với $\bar{Y}$ =125.635)**	**$p_i = P_i - \bar{P}$ (Với $\bar{P}$ =103.887)**	**$p_i \times y_i$**	**p_i^2**	**y_i^2**	**P_i^2**
1	107.919	65.044	-17.716	-38.843	688.135.876	1.508.816.315	313.842.698	4.230.721.936
2	118.866	101.376	-6.769	-2.511	16.999.252	6.307.556	45.814.028	10.277.093.376
3	98.579	124.989	-27.056	21.102	-570.914.281	445.273.942	732.005.819	15.622.250.121
4	122.015	55.249	-3.620	-48.638	176.052.155	2.365.702.208	13.101.548	3.052.452.001
5	152.827	73.775	27.192	-30.112	-818.830.550	906.761.744	739.426.288	5.442.750.625
6	91.259	48.484	-34.376	-55.403	1.904.528.370	3.069.546.133	1.181.682.292	2.350.698.256
7	123.550	138.809	-2.085	34.922	-72.797.602	1.219.512.220	4.345.582	19.267.938.481
8	160.931	50.244	35.296	-53.643	-1.893.421.573	2.877.623.467	1.245.835.425	2.524.459.536
9	98.496	104.300	-27.139	413	-11.195.086	170.169	736.503.939	10.878.490.000
10	108.052	37.852	-17.583	-66.035	1.161.075.916	4.360.685.259	309.148.036	1.432.773.904
11	144.788	66.921	19.153	-36.966	-708.033.647	1.366.521.002	366.852.499	4.478.420.241
12	164.571	166.332	38.936	62.445	2.431.364.241	3.899.317.473	1.516.042.773	27.666.334.224
13	105.564	61.951	-20.071	-41.936	841.690.667	1.758.668.761	402.829.228	3.837.926.401
14	102.568	100.441	-23.067	-3.446	79.498.708	11.878.258	532.068.315	10.088.394.481
15	103.342	39.462	-22.293	-64.425	1.436.211.954	4.150.643.098	496.960.285	1.557.249.444

16	127.030	139.900	1.395	36.013	50.251.645	1.296.901.248	1.947.124	19.572.010.000
17	166.755	171.740	41.120	67.853	2.790.122.153	4.603.963.812	1.690.886.798	29.494.627.600
18	125.343	149.894	-292	46.007	-13.415.779	2.116.599.436	85.034	22.468.211.236
19	121.886	57.386	-3.749	-46.501	174.315.748	2.162.388.093	14.052.047	3.293.152.996
20	134.594	185.105	8.959	81.218	727.659.713	6.596.284.767	80.270.740	34.263.861.025
21	152.937	114.520	27.302	10.633	290.293.117	113.050.378	745.420.715	13.114.830.400
22	109.622	52.933	-16.013	-50.954	815.914.093	2.596.359.526	256.403.553	2.801.902.489
23	149.884	203.500	24.249	99.613	2.415.543.121	9.922.653.175	588.033.106	41.412.250.000
24	98.388	39.334	-27.247	-64.553	1.758.863.372	4.167.152.406	742.377.542	1.547.163.556
25	140.791	95.120	15.156	-8.767	-132.883.454	76.868.791	229.716.277	9.047.814.400
26	101.260	49.200	-24.375	-54.687	1.332.985.900	2.990.720.999	594.121.421	2.420.640.000
27	139.517	113.566	13.882	9.679	134.360.960	93.673.656	192.720.861	12.897.236.356
28	115.236	194.125	-10.399	90.238	-938.344.372	8.142.809.141	108.131.008	37.684.515.625
29	136.749	233.844	11.114	129.957	1.444.387.904	16.888.695.830	123.529.753	54.683.016.336
30	105.067	83.416	-20.568	-20.471	421.049.436	419.081.692	423.026.419	6.958.229.056
31	136.872	183.953	11.237	80.066	899.727.735	6.410.486.716	126.279.023	33.838.706.209
32	117.146	60.457	-8.489	-43.430	368.664.277	1.886.207.014	72.056.433	3.655.048.849
33	163.538	65.065	37.903	-38.822	-1.471.503.937	1.507.185.330	1.436.667.272	4.233.454.225
Tổng	**4.145.942**	**3.428.287**			**15.728.356.029**	**99.938.509.616**	**16.062.183.882**	**456.094.623.385**

Phương sai của $\hat{\beta}_1$ là: Var($\hat{\beta}_1$) = $\frac{\sum P_i^2}{n\sum p_i^2}\sigma^2$

Sai số chuẩn của $\hat{\beta}_1$ là: Se($\hat{\beta}_1$) = $\sqrt{\text{Var}(\hat{\beta}1)}$

Phương sai của $\hat{\beta}_2$ là: Var($\hat{\beta}_2$) = $\frac{\sigma^2}{\sum p_i^2}$

Sai số chuẩn của $\hat{\beta}_2$ là: Se($\hat{\beta}_2$) = $\sqrt{\text{Var}(\hat{\beta}2)}$

Với σ^2 được thay thế $\hat{\sigma}^2 = \frac{\sum y_i^2 - [(\sum p_i y_i)^2/\sum p_i^2]}{n-2} = \frac{16.062.183.882 - \left[\frac{15.728.356.029^2}{99.938.509.616}\right]}{33-2} \approx 438.285.482,4$

từ bảng dữ liệu, ta có:

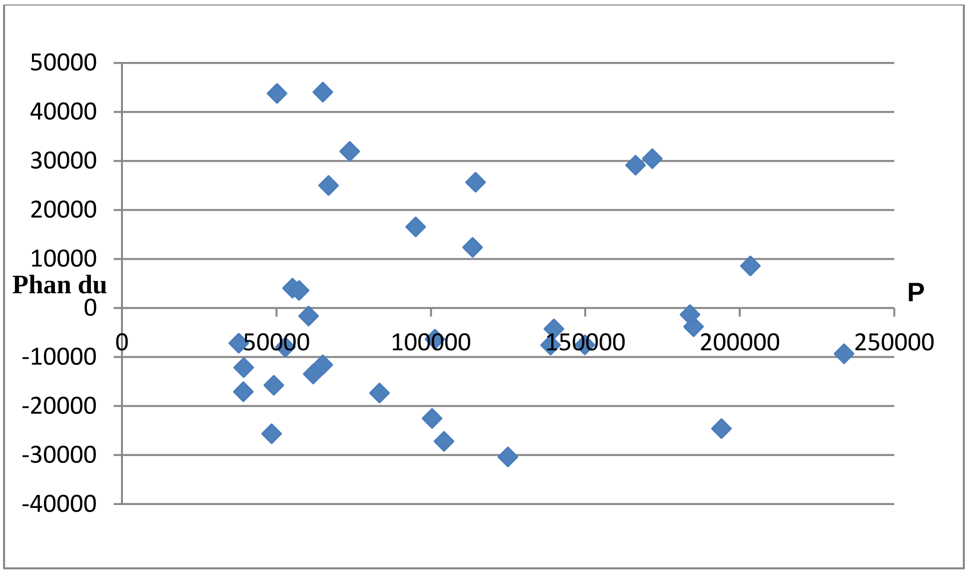
+ Phương sai của $\hat{\beta}_1$ là: Var($\hat{\beta}1$) = $\frac{\sum P_i^2}{n\sum p_i^2}\sigma^2 = \frac{456.094.623.385}{33 \times 99.938.509.616} \times 438.285.482,4 \approx 60.612.923,26$

+ Sai số chuẩn của $\hat{\beta}_1$ là: $Se(\hat{\beta}1) = \sqrt{Var(\hat{\beta}1)} = \sqrt{60.612.923,26} \approx 7.785,430191$

+ Phương sai của $\hat{\beta}_2$ là: $Var(\hat{\beta}_2) = \frac{\sigma^2}{\sum p_i^2} = \frac{438.285.482,4}{99.938.509.616} \approx 0,004385552$

+ Sai số chuẩn của $\hat{\beta}_2$ là: $Se(\hat{\beta}_2) = \sqrt{Var(\hat{\beta}2)} = \sqrt{0,004385552} \approx 0,066223497$

e. Bạn hãy vẽ biểu đồ phân tán giữa phần dư của mô hình ở câu d(trục tung) với biến P(trục hoành)



f. Sử dụng công cụ Tools/Data Analysis/Regression ta tính được các kết quả và viết được các phương trình sau:

Tương quan tuyến tính giữa Y và P

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0,3925677							
R Square	0,1541094							
Adjusted R Square	0,1268226							
Standard Error	20935,269							
Observations	33							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	1	2475333926	2475333926	5,647766	0,023836465			
Residual	31	13586849956	438285482,4					
Total	32	16062183882						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95,0%*	*Upper 95,0%*
Intercept	109284,76	7785,430191	14,03708675	5,62E-15	93406,26949	125163,2	93406,27	125163,248
X Variable 1	0,1573803	0,066223497	2,376502931	0,023836	0,022316623	0,292444	0,022317	0,29244405

Phương trình hồi quy tuyến tính: $\widehat{Y_i} = \widehat{\beta_1} + \widehat{\beta_2}\, P_i = 109.285 + 0,1574 * P_i$

Nhận xét: Ta thấy các ước lượng của hệ số bêta đều có ý nghĩa thống kê ở mức ý nghĩa 5%. Các biến Y và P có mối quan hệ đồng biến, phù hợp với nhận định ban đầu về mối liên hệ giữa Y và P.
Với hệ số β_2= 0,1574 cho ta biết trong các điều kiện khác không đổi nếu số người dân sống xung quanh nhà hàng trong phạm vi bán kính 3 dặm tăng 1 người thì số lượt người được nhà hàng phục vụ trong một năm tăng 0,1574 lượt về mặt trung bình.

Tương quan tuyến tính giữa Y và N

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0,144225							
R Square	0,020801							
Adjusted R	-0,01079							
Standard Er	22524,59							
Observatior	33							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	1	334105420,6	3,34E+08	0,658520878	0,423269919			
Residual	31	15728078461	5,07E+08					
Total	32	16062183882						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	133032	9923,289504	13,40604	1,92166E-14	112793,3048	153270,67	112793,3048	153270,6694
X Variable	-1683,54	2074,622845	-0,81149	0,423269919	-5914,763074	2547,6793	-5914,76307	2547,67927

Phương trình hồi quy tuyến tính: $\widehat{Y_i} = \widehat{\beta_1} + \widehat{\beta_2}\, N_i$ = 133032 -1683,53* N_i

Nhận xét: Ta thấy ước lượng của hệ số bêta không có ý nghĩa thống kê ở mức ý nghĩa 5% do p-value có giá trị lớn, tuy nhiên trên thực tế và các cơ sở lý thuyết đều thấy số đối thủ cạnh tranh trực tiếp có ảnh hưởng đến số lượt khách hàng được phục vụ trong năm của nhà hàng, mối quan hệ này là nghịch biến và phù hợp với nhận định ban đầu về mối liên hệ giữa Y và N.
Với hệ số β_2= -1.683,53 cho ta biết trong các điều kiện khác không đổi nếu số đối thủ cạnh tranh trực tiếp với nhà hàng tăng lên 1 đơn vị thì số lượt người được nhà hàng phục vụ trong một năm sẽ giảm đi khoảng 1.684 lượt về mặt trung bình.

Tương quan tuyến tính giữa Y và I

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0,537022							
R Square	0,288393							
Adjusted R	0,265438							
Standard Er	19201,79							
Observatior	33							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	1	4632215666	4,63E+09	12,56334951	0,001271677			
Residual	31	11429968216	3,69E+08					
Total	32	16062183882						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	77543,48	13973,56728	5,549297	4,44574E-06	49044,19791	106042,75	49044,19791	106042,7544
X Variable	2,339908	0,660155129	3,544482	0,001271677	0,993512424	3,6863029	0,993512424	3,686302939

Phương trình hồi quy tuyến tính: $\widehat{Y_i} = \widehat{\beta_1} + \widehat{\beta_2}$ I_i = 77.643,48+2,3399* I_i

Nhận xét Ta thấy các ước lượng của hệ số bêta đều có ý nghĩa thống kê ở mức ý nghĩa 5%. Các biến Y và I có mối quan hệ đồng biến, phù hợp với nhận định ban đầu về mối liên hệ giữa Y và I. Với hệ số β_2= 2,3399 cho ta biết trong các điều kiện khác không đổi nếu thu nhập trung bình một hộ của những hộ ở quanh nhà hàng trong phạm vi bán kính 3 dặm tăng lên 1 USD thì số lượt người được nhà hàng phục vụ trong một năm sẽ tăng lên khoảng 2,3399 lượt về mặt trung bình.