CHƯƠNG 2

MÔ HÌNH HỒI QUI HAI BIẾN

MÔ HÌNH

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

- $\beta_1 = (Y_i - u_i / X_i = 0) =$ E(Y/X$_i$ =0)

$\Rightarrow$ β_1 cho biết giá trị trung bình của biến phụ thuộc khi giá trị của biến độc lập bằng 0.

- $$\beta_2 = \frac{\mathrm{dE}(\mathrm{Y}/\mathrm{X})}{\mathrm{dX}}$$

$\Rightarrow$ β_2 cho biết khi X tăng lên 1 đơn vị thì giá trị trung bình của biến phụ thuộc thay đổi (tăng, giảm) β_2 đơn vị.

I. PHƯƠNG PHÁP BÌNH PHƯƠNG NHỎ NHẤT (OLS: ordinary least squares)

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

$$e_i = Y_i - \hat{Y}_i$$

$$Y_i = \hat{Y}_i + e_i$$

$\sum e_i^2 \quad \Rightarrow$ min (bình phương nhỏ nhất)

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2 \quad \Rightarrow \text{min}$$

I.1. Các ước lượng OLS

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

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

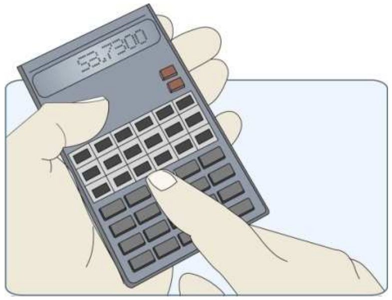
I.1. Các ước lượng OLS

$$\boxed{\hat{\beta}_2 = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}}$$

trong đó:

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \qquad \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$y_i = Y_i - \overline{Y} \qquad x_i = X_i - \overline{X}$$



Ví dụ 2.1

- Giả sử có 5 quan sát về tỉ suất lợi nhuận của công ty máy tính Apple (Y %) và tỉ suất lợi nhuận bình quân của 500 công ty lớn khác ở Mỹ (X%) như sau

X	10	-5	10	-5	-10
Y	20	-5	25	-30	-10

- Tính $\sum X_i, \sum Y_i, \sum X_i Y_i, \sum X_i^2$

- và ước lượng của các hệ số chặn, hệ số góc trong hồi qui $Y_i = \beta_1 + \beta_2 X_i + u_i$

Ví dụ 2.2

$n = 14$; $\sum Y_i = 4444{,}9$; $\sum X_i = 26753$;

$\sum x_i^2 = 55462515$; $\sum x_i y_i = 9095985{,}5$

Tính các ước lượng của β_1, β_2?

Ví dụ 2.3

Cho hàm hồi qui mẫu (SRF) $Y_i = \hat{\beta}_2 X_i + e_i$ (không có hệ số chặn). Viết phương trình biểu diễn $\sum e_i^2$ theo X_i, Y_i, từ đó, rút ra công thức cho ước lượng OLS.


Ví dụ 2.4,
câu 1-2

KỲ VỌNG

- Định nghĩa

  + $E(X) = \Sigma x_i f(x_i)$

- Các tính chất

  + $E(X+c) = E(X) + c$, c là hằng số

  + $E(X+Y) = E(X) + E(Y)$

  + $E(cX) = cE(X)$, c là hằng số

  + $E(XY) = E(X)E(Y)$, X và Y độc lập

Phương sai

- Định nghĩa: $Var(X) = E(X-E(X))^2$
- Tính chất:

  + $Var(cX) = c^2Var(X)$ (c là hằng số)

  + $Var(c+X) = Var(X)$

  + $Var(X+Y) = Var(X)+Var(Y)$, X và Y độc lập

  + $Var(X \pm Y) = Var(X) + Var(Y) \pm 2cov(X, Y)$, X và Y không độc lập

I.2. Các giả thiết của phương pháp ước lượng OLS

1. Các biến giải thích là phi ngẫu nhiên, tức là giá trị của chúng đã được xác định.

2. Kỳ vọng của các yếu tố ngẫu nhiên u bằng 0, $E(u|X_i) = 0$

3. Phương sai của u_i thuần nhất (bằng nhau)

$var(u|X_i) = \sigma^2$ (với $\forall i$)

4. Không có tự tương quan giữa các yếu tố ngẫu nhiên $Cov(u_i, u_j|X_i, X_j) = 0$ (với $\forall i \neq j$)

5. u và X không tương quan với nhau

$Cov(u_i, X_i) = 0$

I.3. Một số tính chất của hàm hồi qui mẫu

1. Đường hồi qui mẫu đi qua trung bình mẫu

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

I.3. Một số tính chất của hàm hồi qui mẫu

2. Tổng các phần dư bằng 0

$$\sum_{i=1}^{n} e_i = 0 \quad \text{hay} \quad \overline{e} = 0$$

I.3. Một số tính chất của hàm hồi qui mẫu

3. Với biến phụ thuộc, giá trị trung bình mẫu bằng giá trị trung bình tổng thể

$$\overline{\hat{Y}} = \overline{Y}$$

I.3. Một số tính chất của hàm hồi qui mẫu

4. Phần dư trực giao với X_i

$$\sum_{i=1}^{n} e_i X_i = 0 \quad \text{hay}$$

$$\sum_{i=1}^{n} e_i x_i = \sum_{i=1}^{n} e_i \left(X_i - \overline{X} \right) = \sum_{i=1}^{n} e_i X_i = 0$$

I.3. Một số tính chất của hàm hồi qui mẫu

5. Phần dư trực giao với giá trị dự báo $\hat{Y}_i$

$$\sum_{i=1}^{n} \hat{Y}_i e_i = 0$$

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

I.4. **Định lý Gauss-Markov:**

Với <u>5 giả thiết đã nêu</u> của phương pháp bình phương nhỏ nhất, các ước lượng nhận được từ phương pháp OLS là các ước lượng <u>tuyến tính</u>, <u>không chệnh</u> và <u>có phương sai nhỏ nhất</u> (BLUE: best linear unbias estimator)

Ước lượng OLS là tuyến tính

$$\Rightarrow \hat{\beta}_2 = \sum_{i=1}^{n} k_i Y_i$$

$$k_i = \frac{x_i}{\sum_{i=1}^{n} x_i^2}$$

Ước lượng OLS không chệch

$$E(\hat{\beta}_2) = \beta_2$$

$$\hat{\beta}_2 = \frac{\sum_{i=1}^{n} y_i x_i}{\sum_{i=1}^{n} x_i^2} = \beta_2 + \frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}$$

Ước lượng OLS
có phương sai nhỏ nhất

$$var(\widehat{\beta}_2) = \sigma^2_{\widehat{\beta}_2} = \frac{\sigma^2}{\sum_{i=1}^{n} x_i^2}$$

Phương sai của ước lượng phụ thuộc vào: phương sai sai số σ^2, số quan sát n,

độ biến thiên của X ($\sum_{i=1}^{n} x_i^2$)

II. Độ chính xác của ước lượng OLS

$$\sigma_{\hat{\beta}_2} = \sqrt{var(\hat{\beta}_2)} = \frac{\sigma}{\sqrt{\sum_{i=1}^{n} x_i^2}}$$

$$var(\hat{\beta}_1) = \frac{\sum_{i=1}^{n} X_i^2}{n\sum_{i=1}^{n} x_i^2}\sigma^2 \quad \Rightarrow \quad \sigma_{\hat{\beta}_1} = \sqrt{\frac{\sum_{i=1}^{n} X_i^2}{n\sum_{i=1}^{n} x_i^2}}.\sigma$$

II. Độ chính xác của ước lượng OLS

Ước lượng của phương sai sai số

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} e_i^2}{n-2}$$

$\hat{\sigma}^2$ được gọi là ước lượng OLS của σ^2 và là ước lượng không chệnh.

n – 2 là số bậc tự do (df: *degree of freedom*)

- Do không có được σ^2, sử dụng $\hat{\sigma}^2$ thay cho σ^2

$$\hat{\sigma}^2_{\hat{\beta}_2} = \frac{\hat{\sigma}^2}{\sum_{i=1}^{n} x_i^2} \quad \Rightarrow \quad se(\hat{\beta}_2) = \frac{\hat{\sigma}}{\sqrt{\sum_{i=1}^{n} x_i^2}}$$

$$\hat{\sigma}^2_{\hat{\beta}_1} = \frac{\sum_{i=1}^{n} X_i^2}{n\sum_{i=1}^{n} x_i^2}\hat{\sigma}^2 \quad \Rightarrow \quad se(\hat{\beta}_1) = \sqrt{\frac{\sum_{i=1}^{n} X_i^2}{n\sum_{i=1}^{n} x_i^2}}.\hat{\sigma}$$


Ví dụ 2.4,
câu 3

III. Giả thiết về phân phối chuẩn của U_i

- Nhiễu ngẫu nhiên u có phân bố chuẩn với kì vọng bằng 0 và phương sai bằng σ^2.

  $u \sim N(0; \sigma^2)$

- Giả thiết này được coi là giả thiết thứ 6 của phương pháp OLS

Các giả thiết của phương pháp ước lượng OLS

1. Các biến giải thích là phi ngẫu nhiên, tức là giá trị của chúng đã được xác định.

2. $E(u_i|X_i) = 0$

3. $var(u_i|X_i) = \sigma^2$

4. $Cov(u_i, u_j|X_i, X_j) = 0$

5. $Cov(u, X) = 0$

6. $u \sim N(0; \sigma^2)$

III. Giả thiết về phân phối chuẩn của U_i

- Các kết quả trước đây không yêu cầu U_i có phân phối chuẩn
- Nhiều hiện tượng kinh tế, xã hội có phân bố chuẩn
- Giả thiết được đánh giá bằng định lý giới hạn trung tâm
- Giả thiết có thể kiểm định được bằng kiểm định Jarque-Bera

Tính chất của biến ngẫu nhiên phân phối chuẩn (Bnnppc)

- Tổ hợp tuyến tính của một Bnnppc là một Bnnppc
- Các Bnnppc không tương quan với nhau thì độc lập với nhau
- 95% diện tích của ppc nằm trong khoảng [-1,96; 1,96]
- Bnnppc có trung bình bằng 0 và phương sai bằng 1 được gọi là biến chuẩn hoá

Các phân phối liên quan tới phân phối chuẩn

- Phân phối Khi bình phương

$$U_i \sim N(0,1) \Rightarrow \sum_{i=1}^{n} U_i^2 \sim \chi^2(n)$$

- Phân phối Student

$$U \sim N(0,1)$$

$$V \sim \chi^2_{(n)}$$

$$\Rightarrow \quad T = \frac{U}{\sqrt{\frac{V}{n}}} \sim T(n)$$

Các phân phối liên quan tới phân phối chuẩn

- Phân phối Fisher

$$V_1 \sim \chi^2_{(n1)} \qquad \Rightarrow \qquad F = \frac{\frac{V_1}{n_1}}{\frac{V_2}{n_2}} \sim F(n_1; n_2)$$

$$V_2 \sim \chi^2_{(n2)}$$

IV. Các phân bố xác suất

$$U \sim N(0, \sigma^2) \Rightarrow \widehat{\beta}_2 \sim N(\beta_2, \sigma^2_{\widehat{\beta}_2})$$

$$\widehat{\beta}_1 \sim N(\beta_1, \sigma^2_{\widehat{\beta}_1})$$

$$\frac{\hat{\beta}_i - \beta_i}{\sigma_{\hat{\beta}_i}} \sim N(0,1)$$

$$Y \sim N(\beta_1 + \beta_2 X_i, \sigma^2)$$

IV. Các phân bố xác suất

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

$$t_i = \frac{\hat{\beta}_i - \beta_i}{se\left(\hat{\beta}_i\right)} \sim T(n-2)$$

V. Phân tích hệ số mô hình hồi qui

1.Ước lượng khoảng tin cậy cho hệ số hồi qui

- Khoảng tin cậy đối xứng

$$\left[\hat{\beta}_i - Se\left(\hat{\beta}_i\right).t_{\alpha/2}{}^{(n-2)}; \hat{\beta}_i + Se\left(\hat{\beta}_i\right).t_{\alpha/2}{}^{(n-2)}\right]$$

- Khoảng tin cậy tối đa

$$\left[-\infty; \hat{\beta}_i + Se\left(\hat{\beta}_i\right).t_{\alpha}{}^{(n-2)}\right]$$

- Khoảng tin cậy tối thiểu

$$\left[\hat{\beta}_i - Se\left(\hat{\beta}_i\right).t_{\alpha}{}^{(n-2)}; +\infty\right]$$


Ví dụ 2.4,
câu 4

V. Phân tích hệ số mô hình hồi qui

2. Kiểm định giả thiết

$$t_i = \frac{\hat{\beta}_i - \beta_i}{se\left(\hat{\beta}_i\right)} \sim T(n-2)$$

Giả thiết: Ho: $\beta_i = \beta_i^*$ (i=1,2)

- *Trường hợp1:* H_1: $\beta_i \neq \beta_i^*$

$|t_i| > T_{\alpha/2}^{(n-2)} \Rightarrow$ Bác bỏ giả thiết H_0

$|t_i| \leq T_{\alpha/2}^{(n-2)} \Rightarrow$ Không đủ cở sở bác bỏ H_0

2. Kiểm định giả thiết

- *Trường hợp 2:* H_1: $\beta_i > \beta_i^*$

$t_i > T_\alpha^{(n-2)} \Rightarrow$ Bác bỏ giả thiết H_0

$t_i \leq T_\alpha^{(n-2)} \Rightarrow$ Không đủ cở sở bác bỏ H_0

- *Trường hợp 3:* H_1: $\beta_i < \beta_i^*$

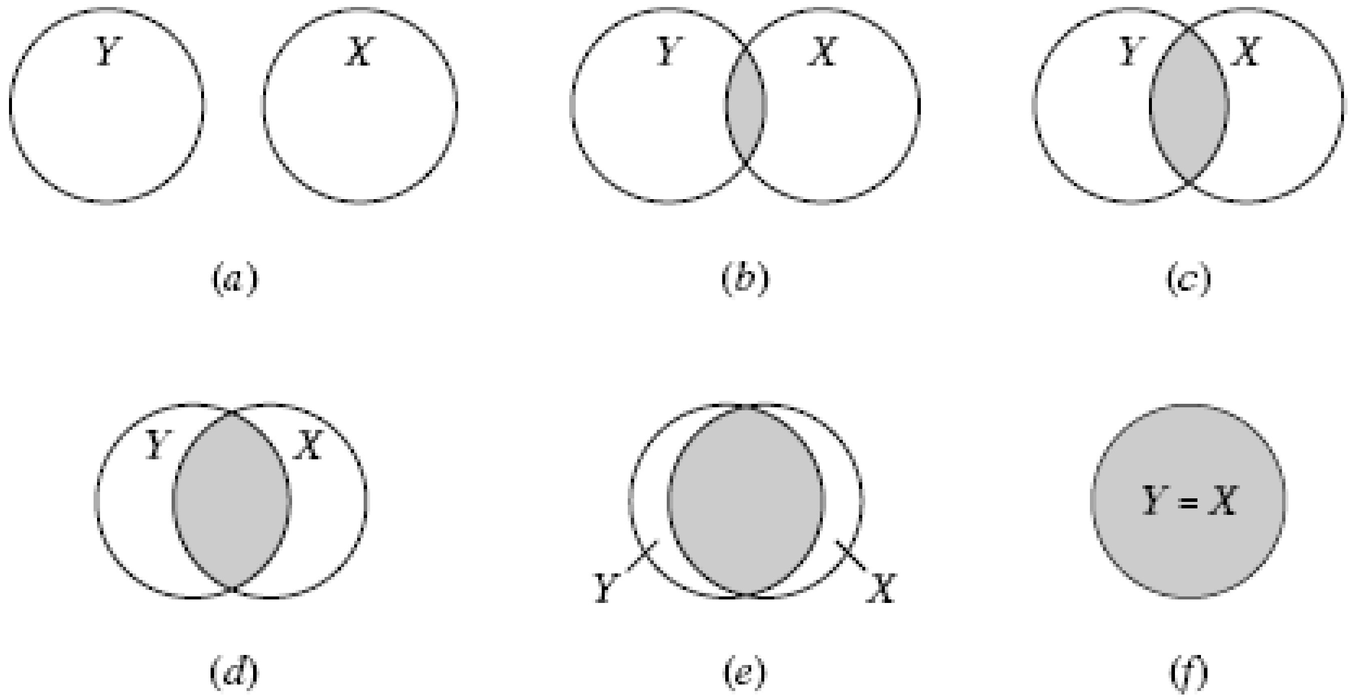
$t_i < -T_\alpha^{(n-2)} \Rightarrow$ Bác bỏ giả thiết H_0

$t_i \geq -T_\alpha^{(n-2)} \Rightarrow$ Không đủ cở sở bác bỏ H_0


Ví dụ 2.4,
câu 5-6

V. Sự phù hợp của hàm hồi qui

1. Hệ số xác định, r^2



(a): $r^2 = 0$ (f): $r^2 = 1$

(b), (c), (d), (e): có r^2 tăng dần và ($0 < r^2 < 1$)

1. Hệ số r^2

$$\sum_{i=1}^{n} y_i^2 = \sum_{i=1}^{n} \hat{y}_i^2 + \sum_{i=1}^{n} e_i^2$$

$$TSS = \sum y_i^2 = \sum (Y_i - \bar{Y})^2$$

$$ESS = \sum \hat{y}_i^2 = \sum (\hat{Y}_i - \bar{Y})^2$$

$$RSS = \sum e_i^2$$

TSS = ESS + RSS

1. Hệ số r^2

- TSS có số bậc tự do là n -1
- ESS có số bậc tự do là 1
- RSS có số bậc tự do là n-2

1. Hệ số r^2

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

- r^2 cho biết có bao nhiêu % sự thay đổi của biến phụ thuộc Y trong mô hình được giải thích bởi biến độc lập X.

$\Rightarrow$ r^2 được sử dụng để đo sự phù hợp của hàm hồi qui

- $1 \geq r^2 \geq 0$

1. Hệ số r^2

$$r^2 = \frac{\text{ESS}}{\text{TSS}} = \frac{\sum_{i=1}^{n} \hat{y}_i^2}{\sum_{i=1}^{n} y_i^2} = \frac{\hat{\beta}_2^2 \sum_{i=1}^{n} x_i^2}{\sum_{i=1}^{n} y_i^2}$$

$$= \frac{\left(\sum_{i=1}^{n} x_i y_i\right)^2}{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i^2} = r_{\text{yx}}^2 = r_{xy}^2$$

Hệ số tương quan mẫu $r = \pm\sqrt{r^2}$

- $-1 \le r \le 1$: r âm hoặc dương tuỳ thuộc vào quan hệ giữa X và Y là cùng chiều hay ngược chiều
- r có tính đối xứng, r(X, Y) = r(Y, X)
- X* = aX + c; Y* = bY + d; với (a, b > 0; c,d: constant) $\Rightarrow$ r(X, Y) = r(X*, Y*)
- X, Y độc lập $\Rightarrow$ r = 0, nhưng r = 0 không thể suy ra X, Y độc lập
- r đo quan hệ phụ thuộc tuyến tính, không đo quan hệ phi tuyến, không phản ánh quan hệ nhân - quả

2. Kiểm định sự phù hợp của hàm hồi qui

- Giả thiết: Ho: $r^2 = 0$ hay $\beta_2 = 0$

  H_1: $r^2 \neq 0$ hay $\beta_2 \neq 0$

$$F = \frac{\hat{\beta}_2^2 \sum x_i^2}{\sum e_i^2 \Big/ (n-2)} = \frac{\mathrm{ESS}/1}{\mathrm{RSS}/(\mathrm{n}-2)} \sim F(1, n-2)$$

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

2. Kiểm định sự phù hợp của hàm hồi qui

- Nếu $F > F\alpha(1,n-2)$ bác bỏ giả thiết H_0 với mức ý nghĩa α.
- Nếu $F < F\alpha(1,n-2)$ không có cơ sở bác bỏ giả thiết H_0 với mức ý nghĩa α.


Ví dụ 2.4,
câu 7-8

Phân tích phương sai (ANOVA) (*analysis of variance*)

Nguồn biến thiên	Tổng bình phương	Bậc tự do	MSS (Mean of Sum of Square)
Do hồi qui (ESS)	$\hat{\beta}_2^2 \sum_{i=1}^{n} x_i^2 = \sum \hat{y}_i^2$	1	$\hat{\beta}_2^2 \sum_{i=1}^{n} x_i^2 / 1$
Do phần dư (RSS)	$\sum_{i=1}^{n} e_i^2$	n-2	$\sum_{i=1}^{n} e_i^2 / (n-2)$
TSS	$\sum_{i=1}^{n} y_i^2$	n-1	$\sum_{i=1}^{n} y_i^2$

VI. Dự báo

1. Dự báo giá trị trung bình

$\hat{Y}$ là ước lượng điểm của E(Y/X_0) và là một ước lượng BLUE, với X = X_0.

Tìm khoảng tin cậy của $\hat{Y}$

$$E(\hat{Y}_0) = \beta_1 + \beta_2 X_0 = E(Y/X_0)$$

$$Var(\hat{Y}_0) = var(\hat{\beta}_1 + \hat{\beta}_2 X_0)$$

$$= var(\hat{\beta}_1) + X_0^2 var(\hat{\beta}_2) + 2X_0 \operatorname{cov}(\hat{\beta}_1, \hat{\beta}_2)$$

VI. Dự báo _ giá trị trung bình

$$Var(\hat{Y}_0) = \sigma^2 \left(\frac{1}{n} + \frac{\left(X_0 - \bar{X}\right)^2}{\sum_{i=1}^{n} x_i^2} \right)$$

- σ^2 chưa biết $\Rightarrow$ sử dụng $\hat{\sigma}^2$

$$se(\hat{Y}_0) = \hat{\sigma} \sqrt{\frac{1}{n} + \frac{\left(X_0 - \bar{X}\right)^2}{\sum_{i=1}^{n} x_i^2}}$$

VI. Dự báo_ giá trị trung bình

$$t = \frac{\hat{Y}_0 - E\left(Y / X\right)}{se\left(\hat{Y}\right)} \sim T(n-2)$$

Mức ý nghĩa α, khoảng tin cậy của $E(Y/X_0)$:

$$\left[\hat{Y}_0 - t_{\alpha/2}{}^{(n-2)}.Se\left(\hat{Y}_0\right); \hat{Y}_0 + t_{\alpha/2}{}^{(n-2)}.Se\left(\hat{Y}_0\right)\right]$$

VI. Dự báo

2. Dự báo giá trị cá biệt, Y_0

Dự báo $Y_0 = \beta_1 + \beta_2 X_0 + u_0$ từ $\hat{Y}_0 = \hat{\beta}_1 + \hat{\beta}_2 X_0$

$$E\left(Y_0 - \hat{Y}_0\right)^2 = \operatorname{var}(\hat{\beta}_1) + X_0^2 \operatorname{var}(\hat{\beta}_2) + 2X_0 \operatorname{cov}(\hat{\beta}_1, \hat{\beta}_2) + \operatorname{var}(u_0)$$

$$Var(Y_0 - \hat{Y}_0) = \sigma^2 \left(1 + \frac{1}{n} + \frac{\left(X_0 - \bar{X}\right)^2}{\sum_{i=1}^{n} x_i^2} \right)$$

VI. Dự báo _ giá trị cá biệt

- σ^2 chưa biết $\Rightarrow$ sử dụng $\hat{\sigma}^2$

$$se(Y_0 - \hat{Y}_0) = \hat{\sigma}\sqrt{1 + \frac{1}{n} + \frac{\left(X_0 - \bar{X}\right)^2}{\sum_{i=1}^{n} x_i^2}}$$

- Khoảng tin cậy (1-α) của Y_0

$$\left[\hat{Y}_0 - t_{\alpha/2}{}^{(n-2)}.Se\left(Y_0 - \hat{Y}_0\right); \hat{Y}_0 + t_{\alpha/2}{}^{(n-2)}.Se\left(Y_0 - \hat{Y}_0\right)\right]$$

Dependent Variable: Y				Y là biến phụ thuộc
Method: Least Squares				Phương pháp OLS
Included observations: 10				Bao gồm 10 quan sát

(Biến) Variable	*Hệ số* Coefficient	*Độ lệch chuẩn* Std. Error	*Thống kê T* t-Statistic	Prob.
C	24.45455	6.413817	3.812791	0.0051
X	0.509091	0.035743	14.24317	0.0000

R-squared	0.962062	Mean dependent var	111.0000
Adjusted R-squared	0.957319	S.D. dependent var	31.42893
S.E. of regression	6.493003	F-statistic	202.8679
Sum squared resid	337.2727	Prob(F-statistic)	0.000001
Log likelihood	-31.78092	Durbin-Watson stat	2.680127

Đinh nghĩa P-value

"Lý thuyết xác suất thống kê", NXB Giáo Dục 2002

Trường Đại học Kinh Tế Quốc Dân

TS. Nguyễn Cao Văn (chủ biên)

TS. Trần Thái Ninh

- Trang 470
- H_1: $\mu > \mu$ thì P – value = $P(U > U_{qs})$
- H_1: $\mu < \mu$ thì P – value = $P(U < U_{qs})$
- H_1: $\mu \neq \mu$ thì P – value = $P(U > |U_{qs}|)$

Nguyên tắc kiểm định bằng Prob

- Kiểm định 2 phía:

  Prob < $\alpha \Rightarrow$ Bác bỏ giả thiết H_0

  Prob > $\alpha \Rightarrow$ Không đủ cơ sở để bác bỏ H_0

- Kiểm định 1 phía:

  Prob/2 < $\alpha \Rightarrow$ Bác bỏ giả thiết H_0

  Prob/2 > $\alpha \Rightarrow$ Không đủ cơ sở để bác bỏ H_0

Một số trường hợp đặc biệt của mô hình hồi qui đơn

Hồi qui qua gốc tọa độ - Hồi qui không có hệ số chặn

- Y = βX + u

$$\hat{\beta}_2 = \frac{\sum X_i Y_i}{\sum X_i^2}$$

$$var(\hat{\beta}_2) = \frac{\sigma^2}{\sum X_i^2}$$

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{n-1}$$

- Y = $\beta_1 + \beta_2$X + u

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

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

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{n-2}$$

Hồi qui qua gốc tọa độ

- r^2 thu được từ mô hình không có hệ số chặn có thể âm và ta không thể trực tiếp so sánh với r^2 trong mô hình có hệ số chặn
- Chỉ nên dùng mô hình hồi qui không có hệ số chặn khi có lý do đặc biệt.

Độ lớn và đơn vị của biến số

Gross Private Domestic Investment and GDP, USA

YEAR	GPDI(bl)	GPDI(ml)	GDP(bl)	GDP(ml)
	Y	Y*	X	X*
1988	828.2	828200	5865.2	5865200
1989	863.5	863500	6062	6062000
1990	815	815000	6136.3	6136300
1991	738.1	738100	6079.4	6079400
1992	790.4	790400	6244.4	6244400
1993	863.6	863600	6389.6	6389600
1994	975.7	975700	6610.7	6610700
1995	996.1	996100	6761.6	6761600
1996	1084.1	1084100	6994.8	6994800
1997	1206.4	1206400	7269.8	7269800

Độ lớn và đơn vị của biến số

$Y^* = 1000\ Y$ $X^* = 1000X$

Hồi qui: $Y = \beta_1 + \beta_2 X + u$

$Y^* = \beta^*_1 + \beta^*_2 X^* + u^*$

Tìm mối quan hệ giữa:

$$\hat{\beta}_1 \leftrightarrow \hat{\beta}_1^*; \quad \hat{\beta}_2 \leftrightarrow \hat{\beta}_2^*$$

$$\mathrm{var}(\hat{\beta}_1) \leftrightarrow \mathrm{var}(\hat{\beta}_1^*); \quad \mathrm{var}(\hat{\beta}_2) \leftrightarrow \mathrm{var}(\hat{\beta}_2^*)$$

$$\hat{\sigma}^2 \leftrightarrow \hat{\sigma}^{*2}; \quad \mathrm{r} \leftrightarrow \mathrm{r}^{*2}$$

Độ lớn và đơn vị của biến số

Tổng quát: Y* = vY X* = wX

$$\hat{\beta}_2^* = \frac{v}{w}\hat{\beta}_2 \qquad \hat{\beta}_1^* = v\hat{\beta}_1$$

$$\hat{\sigma}^{*2} = v^2\hat{\sigma}^2 \qquad var(\hat{\beta}_2^*) = \left(\frac{v}{w}\right)^2 var(\hat{\beta}_2)$$

$$r^2 = r^{*2} \qquad var(\hat{\beta}_1^*) = v^2 var(\hat{\beta}_1)$$