Chương 2

MÔ HÌNH HỒI QUY HAI BIẾN (phần 2)

III. KiỂM ĐỊNH MÔ HÌNH HỒI QUY

1. Các đại lượng ngẫu nhiên

a. Đại lượng ngẫu nhiên U_i

Theo giả thiết của phương pháp OLS, U_i là đại lượng ngẫu nhiên có giá trị trung bình bằng 0 và phương sai không thay đổi

Giả sử Ui ~ $N(0,\sigma^2)$

Khi đó σ^2 được gọi là phương sai của tổng thể , rất khó tính được nên thường được ước lượng bằng phương sai mẫu

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{n-2} = \frac{\sum (Y_i - \hat{Y}_i)^2}{n-2} = \frac{RSS}{n-2}$$

III. KiỂM ĐỊNH MÔ HÌNH HỒI QUY

1. Các đại lượng ngẫu nhiên

a. Đại lượng ngẫu nhiên U_i

Ta có $Y_i = \beta_1 + \beta_2 X_i + U_i$

Vì U_i ~ $N(0,\sigma^2)$

Nên Y_i ~ $N(\beta_1+\beta_2X_i,\sigma^2)$

III. KiỂM ĐỊNH MÔ HÌNH HỒI QUY

1. Các đại lượng ngẫu nhiên

b. Đại lượng ngẫu nhiên $\hat{\beta}_1, \hat{\beta}_2$

Vì sao $\hat{\beta}_1, \hat{\beta}_2$ là các đại lượng ngẫu nhiên ?

Giả sử : $\hat{\beta}_1 \sim N(\beta_1, \sigma^2_{\hat{\beta}_1})$

$\hat{\beta}_2 \sim N(\beta_2, \sigma^2_{\hat{\beta}_2})$

Trong đó $\sigma^2_{\hat{\beta}_1}$ là phương sai của $\hat{\beta}_1$

$\sigma^2_{\hat{\beta}_2}$ là phương sai của $\hat{\beta}_2$

III. KiỂM ĐỊNH MÔ HÌNH HỒI QUY

1. Các đại lượng ngẫu nhiên

Với

$$\sigma^2_{\hat{\beta}_1} = \frac{\sum X_i^2}{n(\sum X_i^2 - n\bar{X}^2)}\sigma^2 \approx \frac{\sum X_i^2}{n(\sum X_i^2 - n\bar{X}^2)}\hat{\sigma}^2$$

$$\sigma^2_{\hat{\beta}_2} = \frac{\sigma^2}{\sum X_i^2 - n\bar{X}^2} \approx \frac{\hat{\sigma}^2}{\sum X_i^2 - n\bar{X}^2}$$

$se(\hat{\beta}_1) = \sqrt{\sigma^2_{\hat{\beta}_1}}$ độ lệch chuẩn của $\hat{\beta}_1$

$se(\hat{\beta}_2) = \sqrt{\sigma^2_{\hat{\beta}_2}}$ độ lệch chuẩn của $\hat{\beta}_2$

III. KiỂM ĐỊNH MÔ HÌNH HỒI QUY

1. Các đại lượng ngẫu nhiên

Vì : $\hat{\beta}_1 \approx N(\beta_1, \sigma^2_{\hat{\beta}_1})$

$\hat{\beta}_2 \approx N(\beta_2, \sigma^2_{\hat{\beta}_2})$

Nên :

$$\frac{\hat{\beta}_1 - \beta_1}{se(\hat{\beta}_1)} \approx N(0,1)$$

$$\frac{\hat{\beta}_2 - \beta_2}{se(\hat{\beta}_2)} \approx N(0,1)$$

Nhưng do σ^2 ước lượng bằng $\hat{\sigma}^2$ dẫn đến

$$\frac{\hat{\beta}_1 - \beta_1}{se(\hat{\beta}_1)} \approx T(n-2)$$

$$\frac{\hat{\beta}_2 - \beta_2}{se(\hat{\beta}_2)} \approx T(n-2)$$

Với *T(n-2)* là phân phối T-Student với bậc tự do *(n-2)*

III. KiỂM ĐỊNH MÔ HÌNH HỒI QUY

2. Các khoảng tin cậy

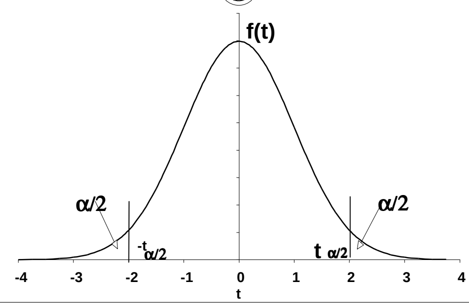
a. *Khoảng tin cậy của β_2*

$$Ta \quad có \quad t = \frac{\hat{\beta}_2 - \beta_2}{se(\hat{\beta}_2)} \approx T(n-2)$$

Giả sử ta muốn xây dựng một khoảng giá trị của β_2 với độ tin cậy (1-α) .

Ví dụ (1-α) = 95% hay 0,95

Đồ thị phân phối của thống kê t



III. KiỂM ĐỊNH MÔ HÌNH HỒI QUY

2. Các khoảng tin cậy

a. *Khoảng tin cậy của β_2*

$$Vì \qquad P\left(-t_{\alpha/2} \le \frac{\hat{\beta}_2 - \beta_2}{se(\hat{\beta}_2)} \le t_{\alpha/2}\right) = 1 - \alpha$$

Nên khoảng tin cậy của β_2 với độ tin cậy 1-α là

$$\boxed{\left(\hat{\beta}_2 - t_{\frac{\alpha}{2}} \times se(\hat{\beta}_2);\ \hat{\beta}_2 + t_{\frac{\alpha}{2}} \times se(\hat{\beta}_2)\right)}$$

Với $t_{\alpha/2}$ *có được khi tra bảng* ***t-Student*** *với bậc tự do* ***(n-2),*** *mức ý nghĩa* ***α/2***

III. KiỂM ĐỊNH MÔ HÌNH HỒI QUY

2. Các khoảng tin cậy

b. *Khoảng tin cậy của β_1*

$$Vì \qquad t = \frac{\hat{\beta}_1 - \beta_1}{se(\hat{\beta}_1)} \approx T(n-2)$$

Lập luận tương tự, khoảng tin cậy của β_1 với độ tin cậy 1-α là

$$\boxed{\left(\hat{\beta}_1 - t_{\frac{\alpha}{2}} \times se(\hat{\beta}_1);\ \hat{\beta}_1 + t_{\frac{\alpha}{2}} \times se(\hat{\beta}_1)\right)}$$

Giải thích ý nghĩa của độ tin cậy (1- α), ví dụ (1- α) =95%?

III. KiỂM ĐỊNH MÔ HÌNH HỒI QUY

2. Các khoảng tin cậy

c. *Khoảng tin cậy của σ^2*

Vì $\hat{\sigma}^2$ là ước lượng của σ^2 và người ta chứng minh được rằng

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

Nên khoảng tin cậy của σ^2 với độ tin cậy 1-α là

$$\boxed{\left(\frac{(n-2).\hat{\sigma}^2}{\chi^2_{\alpha/2}};\ \frac{(n-2).\hat{\sigma}^2}{\chi^2_{1-\alpha/2}}\right)}$$

Với $\chi^2_{\alpha/2}$ có được khi tra bảng χ^2 với bậc tự do (n-2), mức ý nghĩa α/2

Ví dụ áp dụng

Từ số liệu đã cho của ví dụ trước , yêu cầu tính khoảng tin cậy của β_1, β_2 và σ^2 với độ tin cậy 95%

III. KiỂM ĐỊNH MÔ HÌNH HỒI QUY

Nhắc lại về giả thiết H_0

Trong thống kê, giả thiết phát biểu cần được kiểm định được gọi là ***giả thiết không*** (ký hiệu : H_0). ***Giả thiết đối*** được ký hiệu là giả thiết H_1

	Báo bỏ H_0	Chấp nhận H_0
H_0 sai	*Đúng*	*Sai lầm loại II*
H_0 đúng	*Sai lầm loại I*	*Đúng*

Người ta thường đặt giả thiết H_0 sao cho sai lầm loại I là nghiêm trọng (nguy hiểm) hơn sai lầm loại II

III. KiỂM ĐỊNH MÔ HÌNH HỒI QUY

Đặt α là khả năng mắc sai lầm loại I

$\Rightarrow$ α là mức ý nghĩa của kiểm định

$\Rightarrow$ 1- α là độ tin cậy của kiểm định

Chú ý

- Khi nói "chấp nhận giả thiết H_0", không có nghĩa H_0 đúng.
- Lựa chọn mức ý nghĩa α : α có thể tùy chọn, thường người ta chọn mức 1%, 5%, hoặc 10%.

III. KiỂM ĐỊNH MÔ HÌNH HỒI QUY

Các giả thiết cần kiểm định gồm

- *Các giả thiết về hệ số hồi quy*
- *Các giả thiết về phương sai của U_i*
- *Các giả thiết về sự phù hợp của mô hình*

Các loại giả thiết

- *Giả thiết 2 phía , phía trái và phía phải*

Các cách kiểm định cơ bản :

- o *Phương pháp khoảng tin cậy*
- o *Phương pháp giá trị tới hạn*
- o *Phương pháp p-value (dùng máy vi tính)*

III. KiỂM ĐỊNH MÔ HÌNH HỒI QUY

3. Kiểm định giả thiết về hệ số hồi quy

a. *Kiểm định giả thiết về β_2*

Giả thiết 2 phía $H_o:\beta_2 = \beta_o$, $H_1:\beta_2 \neq \beta_o$ — độ tin cậy là *1-α*

Giả thiết phía trái $H_o:\beta_2 = \beta_o$, $H_1:\beta_2 < \beta_o$

Giả thiết phía phải $H_o:\beta_2 = \beta_o$, $H_1:\beta_2 > \beta_o$

III. KiỂM ĐỊNH MÔ HÌNH HỒI QUY

3. Kiểm định giả thiết về hệ số hồi quy

a. *Kiểm định giả thiết về β_2*

Phương pháp khoảng tin cậy

Bước 1 : Lập khoảng tin cậy của β_2

Bước 2 : Nếu β_0 thuộc khoảng tin cậy thì chấp nhận H_0. Nếu β_0 không thuộc khoảng tin cậy thì bác bỏ H_0

III. KiỂM ĐỊNH MÔ HÌNH HỒI QUY

3. Kiểm định giả thiết về hệ số hồi quy

Miền bác bỏ | Miền chấp nhận | Miền bác bỏ

Kiểm định hai phía

Miền bác bỏ | Miền chấp nhận

Kiểm định phía trái

Miền chấp nhận | Miền bác bỏ

Kiểm định phía phải

III. KiỂM ĐỊNH MÔ HÌNH HỒI QUY

2. Kiểm định giả thiết về hệ số hồi quy

a. Kiểm định giả thiết về β_2

Phương pháp giá trị tới hạn (kiểm định t)

Bước 1 : tính giá trị tới hạn $t = \dfrac{\hat{\beta}_2 - \beta_0}{se(\hat{\beta}_2)}$

Bước 2 : tra bảng t-Student với bậc tự do (n-2) tìm $t_{\alpha/2}$

Bước 3 :

Nếu $-t_{\alpha/2} \le t \le t_{\alpha/2}$: chấp nhận giả thiết H_0

Nếu $t < -t_{\alpha/2}$ hoặc $t > t_{\alpha/2}$: bác bỏ giả thiết H_0

SV tự suy luận điều kiện cho kiểm định phía trái và phải

III. KiỂM ĐỊNH MÔ HÌNH HỒI QUY

2. Kiểm định giả thiết về hệ số hồi quy

a. Kiểm định giả thiết về β_2

Phương pháp p-value

Bước 1 : tính giá trị tới hạn $t = \dfrac{\hat{\beta}_2 - \beta_0}{se(\hat{\beta}_2)}$

Bước 2 : Tính p_value = $P(|t| > |t_{\alpha/2}|)$

(tức là khả năng giả thiết H_0 bị bác bỏ)

Bước 3 :

Nếu p_value $\ge \alpha$: chấp nhận giả thiết H_0

Nếu p_value $< \alpha$: bác bỏ giả thiết H_0

III. KiỂM ĐỊNH MÔ HÌNH HỒI QUY

2. Kiểm định giả thiết về hệ số hồi quy

b. Kiểm định giả thiết về β_1

$H_o: \beta_1 = \beta_o$
$H_1: \beta_1 \neq \beta_o$ Với độ tin cậy là *1-α*

Tương tự kiểm định giả thiết về β_2 nhưng giá trị tới hạn lúc này là

$$t = \frac{\hat{\beta}_1 - \beta_0}{se(\hat{\beta}_1)}$$

III. KiỂM ĐỊNH MÔ HÌNH HỒI QUY

2. Kiểm định giả thiết về hệ số hồi quy

c. Kiểm định giả thiết về σ^2

$H_o: \sigma^2 = \sigma_0^2$
$H_1: \sigma^2 \neq \sigma_0^2$ Với độ tin cậy là *1-α*

Bước 1 : Lập khoảng tin cậy của σ^2

Bước 2 :

- Nếu σ_0^2 thuộc khoảng tin cậy thì chấp nhận H_0.
- Nếu σ_0^2 không thuộc khoảng tin cậy thì bác bỏ H_0

Ví dụ áp dụng

Từ số liệu đã cho của ví dụ trước , yêu cầu kiểm định các giả thiết sau

a) $H_o: \beta_2 = 0$
$H_1: \beta_2 \neq 0$ Với độ tin cậy là 95%

b) $H_o: \beta_1 = 0$
$H_1: \beta_1 \neq 0$ Với độ tin cậy là 95%

c) $H_o: \sigma^2 = 16$
$H_1: \sigma^2 \neq 16$ Với độ tin cậy là 95%

III. KiỂM ĐỊNH MÔ HÌNH HỒI QUY

4. Kiểm định sự phù hợp của mô hình

Kiểm định giả thiết

$H_o: R^2 = 0$
$H_1: R^2 \neq 0$ Với độ tin cậy là 1- α

Phương pháp kiểm định F

Bước 1 : tính $F = \dfrac{R^2(n-2)}{(1-R^2)}$

Bước 2 : Tra bảng tìm F(1,n-2), mức ý nghĩa là α

Bước 3 : Nếu F>F(1,n-2) , bác bỏ H_0

Nếu F≤F(1,n-2) , chấp nhận H_0

Câu hỏi

Việc kiểm định giả thiết $H_o: \beta_2 = 0$, $H_1: \beta_2 \neq 0$ *độ tin cậy là (1-α)*

có ý nghĩa như thế nào?

Việc kiểm định giả thiết $H_o: R^2 = 0$, $H_1: R^2 \neq 0$ *độ tin cậy là (1-α)*

có ý nghĩa như thế nào?

Ví dụ áp dụng

Từ số liệu đã cho của ví dụ trước , yêu cầu kiểm định sự phù hợp của mô hình với độ tin cậy 95%

5. Đánh giá kết quả hồi quy

- Dấu của các hệ số hồi qui ước lượng được phù hợp với lý thuyết hay tiên nghiệm không.
- Các hệ số hồi qui ước lượng được có ý nghĩa về mặt thống kê hay không ?
- Mức độ phù hợp của mô hình (R^2) và mô hình có thực sự phù hợp?
- Kiểm tra xem mô hình có thỏa mãn *các giả thiết của mô hình* hồi qui tuyến tính cổ điển hay không.