



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



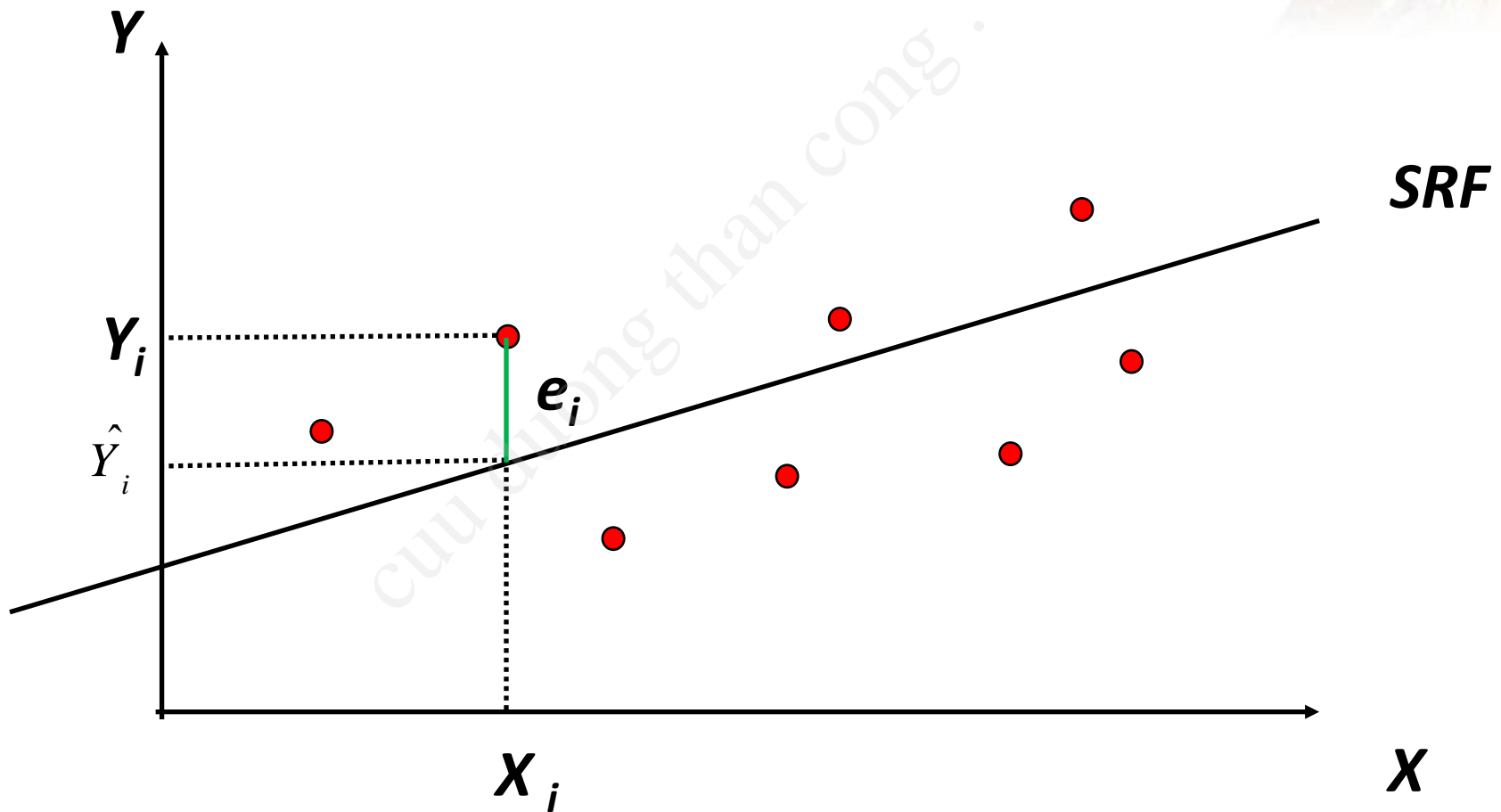
- 1. Ước lượng SRF**
- 2. Các giả thiết cơ bản của phương pháp OLS**
- 3. Độ chính xác của các ước lượng**
- 4. Phân tích hồi qui**
- 5. Kiểm định sự phù hợp của hàm hồi qui**
- 6. m EVIEWS**
- 7. Dự báo**



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



1. ƯỚC LƯỢNG SRF





Chương II – Hồi quy đơn Ước lượng và kiểm định giả thuyết



1. ƯỚC LƯỢNG SRF

Hàm hồi quy mẫu (SRF) : $\hat{Y}_i = \beta_1 + \beta_2 \times X_i$

Mô hình hồi quy mẫu (SRM) : $Y_i = \beta_1 + \beta_2 \times X_i + e_i$

Tiêu chuẩn ước lượng phương pháp OLS:

$$\sum_{i=1}^n e_i^2 \rightarrow \min$$

Tổng bình phương phần dư (***Residual Sum of Squares***)



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



1. Ước lượng SRF

$$Q = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Y_i - \beta_1 - \beta_2 \times X_i)^2 \rightarrow \min$$

Điều kiện để biểu thức trên đạt cực trị:

$$\frac{\partial Q}{\partial \beta_1} = \sum_{i=1}^n 2(Y_i - \beta_1 - \beta_2 \times X_i) \times (-1) = 0$$

$$\frac{\partial Q}{\partial \beta_2} = \sum_{i=1}^n 2(Y_i - \beta_1 - \beta_2 \times X_i) \times (-X_i) = 0$$

Hệ phương trình chuẩn của phương pháp OLS



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



1. Ước lượng SRF

$$\beta_1 = \bar{Y} - \beta_2 \bar{X}$$

$$\beta_2 = \frac{n \sum_{i=1}^n Y_i X_i - \sum_{i=1}^n Y_i \sum_{i=1}^n X_i}{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2}$$

Hoặc:
$$\beta_2 = \frac{\overline{YX} - \bar{X} \cdot \bar{Y}}{\overline{X^2} - (\bar{X})^2}$$

t n:

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

:

$$\beta_2 = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}$$



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



1. Ước lượng SRF

Ví dụ 2.1 (Tr.24 - giáo trình):

Y – lãi suất (đơn vị: %)

X – tỷ lệ lạm phát (đơn vị: %)

$$\hat{\beta}_1 = 2,7418355$$

$$\hat{\beta}_2 = 1,2494067$$

$$SRF : \hat{Y}_i = 2,74 + 1,25X_i$$

(?) Giải thích kết quả ước lượng

(?) Kết quả ước lượng có phù hợp với lý thuyết kinh tế không

Ví dụ: Các ngành công nghiệp Ấn Độ, 1958 (Basic Econometrics – Gujarati)

Ngành công nghiệp (Industry)	Năng suất lao động (VA/L)	Tiền lương thực tế (W)
Lúa mì (Wheat flour)	40.33824	19.33081
Đường (Sugar)	32.4435	17.34319
Sơn (Paints and varnishes)	54.61999	22.55146
Xi măng (Cement)	38.89633	20.84471
Kính và Thủy tinh (Glass/glassware)	25.3328	17.6847
Đồ gốm (Ceramics)	28.26997	19.57983
Gỗ dán (Plywood)	30.90135	16.92345
Dệt bông (Cotton textiles)	27.54442	21.95072
Dệt len (Woolen textiles)	33.32141	20.25902
Hóa chất (Chemicals)	48.53572	21.99687
Nhôm (Aluminum)	41.71664	21.93536
Sắt thép (Iron and Steel)	43.44953	25.16867
Xe đạp (Bicycles)	38.86523	22.25352

$$\left(\frac{VA}{L} \right)_i = \beta_1 + \beta_2 \cdot W_i + U_i \rightarrow \left(\frac{VA}{L} \right)_i = -9,188 + 2,254 \cdot W_i + e_i$$



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



1. Ước lượng SRF

Tính chất của các ước lượng OLS:

- Các β với mẫu cụ thể (mẫu xác định): ước lượng điểm
với mẫu ngẫu nhiên: biến ngẫu nhiên

- Một số tính chất khác: $\sum_{i=1}^n e_i = 0$

$$\bar{Y} = \beta_1 + \beta_2 \bar{X}$$

$$\bar{Y} = \hat{Y}$$

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

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



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



2. Các giả thiết (Assumptions) cơ bản của phương pháp OLS

Các β đại diện cho β cần là các BLUE (best linear unbiased estimator)

- Ước lượng tuyến tính
- Ước lượng không chệch
- Ước lượng hiệu quả (có phương sai nhỏ nhất trong lớp các ước lượng tuyến tính không chệch)

Các ước lượng là BLUE \longleftrightarrow Các giả thiết OLS thỏa mãn



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



2. Các giả thiết (Assumptions) cơ bản của phương pháp OLS

GT 1: Hồi quy tuyến tính (với tham số)

GT 2: Các giá trị của X cố định trong các mẫu lặp lại

GT 3: Kỳ vọng có điều kiện của sai số ngẫu nhiên bằng 0

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

GT 4: Phương sai sai số u_i là đồng đều

$$\text{Var}(u_i | X_i) = \sigma^2$$

GT 5: Không tồn tại TỰ TƯƠNG QUAN giữa các sai số ngẫu nhiên:

$$\text{cov}(u_i, u_j | X_i, X_j) = 0 \quad \text{với} \quad i \neq j$$



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



2. Các giả thiết (Assumptions) cơ bản của phương pháp OLS

GT 6: Không tồn tại tương quan giữa u_i và X_i

$$\text{cov}(u_i, X_i) = E(u_i, X_i) = 0$$

GT 7: Số quan sát n phải lớn hơn số tham số cần ước lượng của mô hình hồi quy

GT 8: Biến độc lập nhận nhiều hơn 1 giá trị

GT 9: Mô hình hồi quy được định dạng đúng

GT 10: Không tồn tại ĐA CỘNG TUYẾN HOÀN HẢO giữa các biến độc lập



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



3. Độ chính xác của các ước lượng

3.1. Độ chính xác của các β

$$\text{var}(\beta_1) = \frac{\sum_{i=1}^n X_i^2}{n \sum_{i=1}^n x_i^2} \sigma^2$$

$$\text{var}(\beta_2) = \frac{\sigma^2}{\sum_{i=1}^n x_i^2}$$

$$\text{cov}(\beta_1, \beta_2) = -\bar{X} \text{var}(\beta_2) = -\bar{X} \left(\frac{\sigma^2}{\sum_{i=1}^n x_i^2} \right)$$

Với:

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

σ^2 : variance of the regression

σ : standard error of the regression



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



3. Độ chính xác của các ước lượng

3.1. Độ chính xác của các β

$$SD(\beta_1) = \sqrt{\text{var}(\beta_1)}$$

$$SD(\beta_2) = \sqrt{\text{var}(\beta_2)}$$

(**Standard deviation**: độ lệch chuẩn)

Khi sử dụng $\sigma^2 \rightarrow \sigma^2 = \text{var}(u|X_i) = \text{var}(Y|X_i)$

Ta có $Se(\beta) = SD(\beta) \rightarrow SD(\beta)$

(**Standard error**: sai số chuẩn)

Phương sai mẫu của biến phụ thuộc càng lớn \rightarrow phương sai và sai số chuẩn của các ước lượng OLS càng lớn \rightarrow các ước lượng OLS càng kém chính xác và ngược lại



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



3. Độ chính xác của các ước lượng

3.1. Độ chính xác của các β

Ví dụ 2.1: $\sum_{i=1}^9 e_i^2 = 20,822895$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^9 e_i^2}{n-2} = \frac{20,822895}{9-2} = 2,9746993$$

$$s.e(\hat{\beta}_1) = 0,6811$$

$$s.e(\hat{\beta}_2) = 0,0388208$$

**Phương sai và sai số chuẩn của các ước lượng càng lớn
→ các ước lượng càng kém chính xác và ngược lại.**



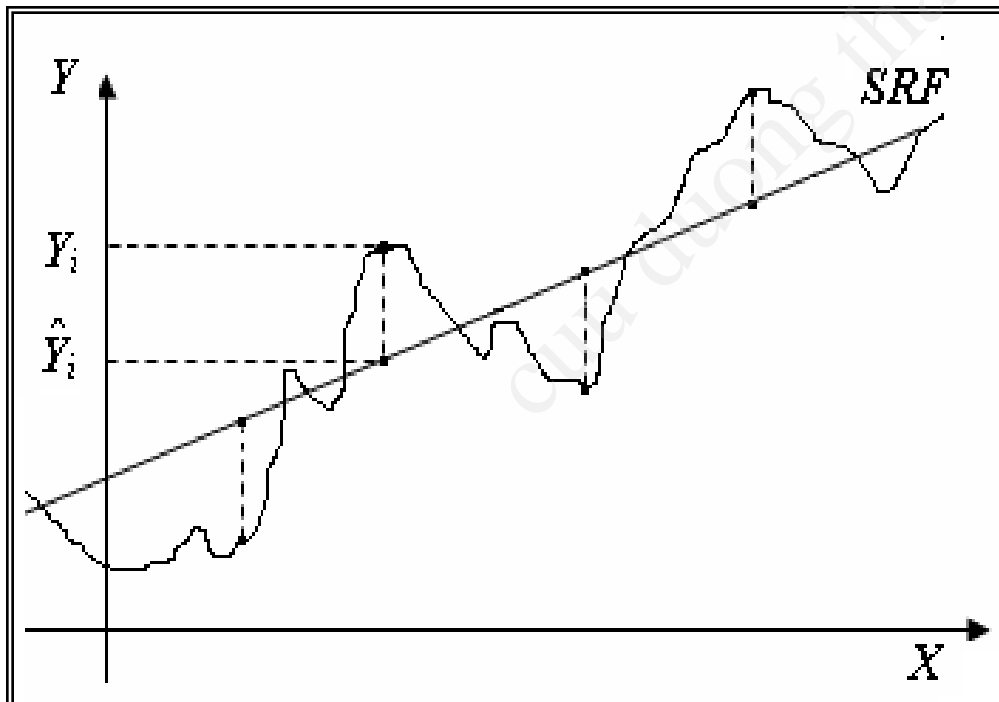
Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



3. Độ chính xác của các ước lượng

3.2. Độ chính xác (độ phù hợp) của SRF:

Toàn bộ sự biến động của biến Y = Sự biến động của biến Y do SRF giải thích + Sự biến động của biến Y do e_i giải thích



$$\text{TSS} = \text{ESS} + \text{RSS}$$

$$\sum_{i=1}^n (Y_i - \bar{Y})^2 = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 + \sum_{i=1}^n e_i^2$$

$$0 \leq R^2 = \frac{\text{ESS}}{\text{TSS}} = 1 - \frac{\text{RSS}}{\text{TSS}} \leq 1$$



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



3. Độ chính xác của các ước lượng

R^2 (R-squared): hệ số xác định của mô hình

Ý nghĩa: Hệ số R^2 cho biết X (hoặc hàm hồi quy mẫu SRF) giải thích được bao nhiêu % sự biến động của Y.

Dependent Variable: IBM			
Method: Least Squares			
Included observations: 120			
Variable	Coefficient	Std. Error	
MARKET	0.453024	0.067675	
C	0.003278	0.004703	
R-squared	0.275235	Mean dependent var	0.009617



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



4. Phân tích hồi qui

Tham số tổng thể
 β_1, β_2 và σ^2

Tham số và các
thông tin khác của mẫu

$$\hat{\beta}_1 - \hat{\beta}_2 - s.e(\hat{\beta}_1) - s.e(\hat{\beta}_2) - \hat{\sigma}^2$$

n: số quan sát
k: số hệ số hồi qui

Dựa trên giả thiết 10 →

$$\hat{\beta}_1 \sim N(\beta_1, \text{var}(\hat{\beta}_1))$$

$$\hat{\beta}_2 \sim N(\beta_2, \text{var}(\hat{\beta}_2))$$

Xây dựng thống kê:

$$T = \frac{\hat{\beta}_j - \beta_j}{s.d(\hat{\beta}_j)} \sim N(0,1)$$

$$T = \frac{\hat{\beta}_j - \beta_j}{s.e(\hat{\beta}_j)} \sim T(n-2)$$



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



4.1. Kiểm định giả thuyết về β_j :

Cặp giả thuyết 1:

$$\begin{cases} H_0 : \beta_j = \beta_j^* \\ H_1 : \beta_j \neq \beta_j^* \end{cases}$$

Tiêu chuẩn kiểm định:

$$T_{qs} = \frac{\beta_j - \beta_j^*}{SE(\beta_j)}$$

Miền bác bỏ H_0 :

$$W_\alpha = T : |T| > T_{\alpha/2}^{(n-2)}$$

T_{qs} thuộc miền bác bỏ $H_0 \rightarrow$ bác bỏ H_0 và ngược lại



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



4.1. Kiểm định giả thuyết về β_j :

Cặp giả thuyết 2:

$$\begin{cases} H_0 : \beta_j = \beta_j^* \\ H_1 : \beta_j > \beta_j^* \end{cases}$$

Tiêu chuẩn kiểm định:

$$T_{qs} = \frac{\beta_j - \beta_j^*}{SE(\beta_j)}$$

Miền bác bỏ H_0 :

$$W_\alpha = T : T > T_\alpha^{(n-2)}$$

T_{qs} thuộc miền bác bỏ $H_0 \rightarrow$ bác bỏ H_0 và ngược lại



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



4.1. Kiểm định giả thuyết về β_j :

Cặp giả thuyết 3:

$$\begin{cases} H_0 : \beta_j = \beta_j^* \\ H_1 : \beta_j < \beta_j^* \end{cases}$$

Tiêu chuẩn kiểm định:

$$T_{qs} = \frac{\beta_j - \beta_j^*}{SE(\beta_j)}$$

Miên bác bỏ H_0 : $W_\alpha = T : T < -T_\alpha^{(n-2)}$

T_{qs} thuộc miền bác bỏ $H_0 \rightarrow$ bác bỏ H_0 và ngược lại



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



Ví dụ 2.2: Y – năng suất (tạ/ha), X – lượng phân bón (tạ/ha)

Dependent Variable: Y		
Method: Least Squares		
Included observations: 10		
Variable	Coefficient	Std. Error
X	1.659722	0.101321
C	27.12500	1.979265

(?) Các hệ số có ý nghĩa thống kê hay không

(?) Không có phân bón, năng suất = 30 tạ/ha

(?) Lượng phân bón có tác động đến năng suất cây trồng hay không

(?) Năng suất tăng hơn 1 tạ/ha khi tăng phân bón 1 tạ/ha



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



Ví dụ: Y – tổng vốn đầu tư (tỉ đồng), X – lãi suất NH (%/năm)

Dependent Variable: Y		
Method: Least Squares		
Included observations: 10		
Variable	Coefficient	Std. Error
X	-9.820896	0.895522
C	94.55224	5.277127

- (?) Lãi suất có thực sự ảnh hưởng đến tổng vốn đầu tư không
- (?) Lãi suất tăng làm tổng vốn đầu tư giảm
- (?) Lãi suất giảm 1% thì vốn đầu tư tăng 9 tỷ. Nhận xét
- (?) Lãi suất tăng 2,5 % thì tổng vốn đầu tư sẽ giảm hơn 20 tỷ.
Nhận xét ý kiến này



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



4.2. Khoảng tin cậy cho β_j :

* Khoảng tin cậy đối xứng:

$$(\beta_j - t_{\alpha/2}^{(n-2)} \times SE(\beta_j); \beta_j + t_{\alpha/2}^{(n-2)} \times SE(\beta_j))$$

* Khoảng tin cậy bên trái (max β_j):

$$(-\infty; \beta_j + t_{\alpha}^{(n-2)} \times SE(\beta_j))$$

* Khoảng tin cậy bên phải (min β_j):

$$(\beta_j - t_{\alpha}^{(n-2)} \times SE(\beta_j); +\infty)$$



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



Ví dụ 2.2: Y – năng suất (tạ/ha), X – lượng phân bón (tạ/ha)

Dependent Variable: Y		
Method: Least Squares		
Included observations: 10		
Variable	Coefficient	Std. Error
X	1.659722	0.101321
C	27.12500	1.979265

- (?) Khi không sử dụng phân bón, năng suất trung bình nằm trong khoảng nào
- (?) Tăng lượng phân bón 1 tạ/ha thì năng suất tăng trong khoảng nào
- (?) Giảm lượng phân bón 1,5 tạ/ha, năng suất giảm tối đa bao nhiêu



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



Ví dụ: Y – tổng vốn đầu tư (tỉ đồng), X – lãi suất NH (%/năm)

Dependent Variable: Y		
Method: Least Squares		
Included observations: 10		
Variable	Coefficient	Std. Error
X	-9.820896	0.895522
C	94.55224	5.277127

(?) Để giảm ảnh hưởng của lạm phát, NH nhà nước dự kiến tăng lãi suất 3%/năm thì tổng vốn đầu tư thay đổi trong khoảng nào

(?) Nếu mức tăng của lãi suất chỉ là 2%/năm thì vốn đầu tư giảm tối đa là bao nhiêu



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



4.3. Khoảng tin cậy cho σ^2 :

* Khoảng tin cậy 2 phía:

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

Hoặc:

$$\left(\frac{RSS}{\chi^2_{\alpha/2}(n-2)}; \frac{RSS}{\chi^2_{1-\alpha/2}(n-2)} \right)$$



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



5. Kiểm định sự phù hợp của hàm hồi quy :

Cặp giả thuyết:

$$\begin{cases} H_0 : R^2 = 0 \\ H_1 : R^2 \neq 0 \end{cases} \Leftrightarrow \begin{cases} H_0 : \beta_2 = 0 \\ H_1 : \beta_2 \neq 0 \end{cases}$$

Tiêu chuẩn kiểm định:

$$F_{qs} = \frac{R^2 / (2 - 1)}{(1 - R^2) / (n - 2)} = F \text{ - statistic}$$

Miền bác bỏ H_0 :

$$W_\alpha = \{ F : F > F_\alpha^{(1, n-2)} \}.$$

Dependent Variable: Y				
Method: Least Squares				
Sample: 1 10				
Included observations: 10				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	27.12500	1.979265	13.70458	0.00000
X	1.659722	0.101321	16.38082	0.00000
R-squared	0.971049	Mean dependent var		57.00000
Adjusted R-squared	0.967430	S.D. dependent var		13.47426
S.E. of regression	2.431706	Akaike info criterion		4.791920
Sum squared resid	47.30556	Schwarz criterion		4.852437
Log likelihood	-21.95960	F-statistic		268.331
Durbin-Watson stat	1.783613	Prob(F-statistic)		0.00000



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



7. Dự báo:

- **Đối tượng của công tác dự báo là:**
 - **Giá trị trung bình của biến phụ thuộc**
 - **Giá trị cá biệt của biến phụ thuộc**
- **Có 2 loại hình dự báo:**
 - **Dự báo bằng ước lượng điểm**
 - **Dự báo bằng khoảng tin cậy**
- **Yêu cầu cho công tác dự báo:**
 - **Độ chính xác và đầy đủ của số liệu**
 - **Thông tin cho trước của biến độc lập trong mô hình**
 - **Khả năng phù hợp với lý thuyết và thực tế của mô hình**



Chương II – Hồi quy đơn Ước lượng và kiểm định giả thuyết



7.1. Dự báo bằng ước lượng điểm:

Với mô hình hồi quy đơn:

$$Y_i = \beta_1 + \beta_2 * X_i + U_i \text{ và giá trị cho trước của } X = X_0$$

Công thức dự báo bằng ước lượng điểm cho giá trị trung bình và giá trị cá biệt của biến phụ thuộc:

$$\hat{Y}_0 = Y_0 = \beta_1 + \beta_2 \times X_0$$

Prediction is difficult especially with regard to the future

- Chinese proverb -



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



7.2. Dự báo bằng khoảng tin cậy:

a/ Cho giá trị trung bình của Y: $E(Y/X_0)$

$$(\hat{Y}_0 - t_{\alpha/2}^{(n-2)} SE(\hat{Y}_0); \hat{Y}_0 + t_{\alpha/2}^{(n-2)} SE(\hat{Y}_0))$$

Với:

$$SE(\hat{Y}_0) = \sqrt{\frac{\sigma^2}{n} + (X_0 - \bar{X})^2 \text{var}(\beta_2)}$$

Trong các công thức trên:

$$\sigma^2 \leftarrow \sigma^2 \quad \bar{X} = \frac{\bar{Y} - \beta_1}{\beta_2} \quad \text{var}(\beta_2) = (SE(\beta_2))^2$$



Chương II – Hồi qui đơn Ước lượng và kiểm định giả thuyết



7.2. Dự báo bằng khoảng tin cậy:

b/ Cho giá trị cá biệt của Y: (Y/X₀)

$$(Y_0 - t_{\alpha/2}^{(n-2)} SE(Y_0); Y_0 + t_{\alpha/2}^{(n-2)} SE(Y_0))$$

Với:

$$SE(Y_0) = \sqrt{\sigma^2 + \frac{\sigma^2}{n} + (X_0 - \bar{X})^2 \text{var}(\beta_2)}$$

Trong các công thức trên:

$$\sigma^2 \leftarrow \sigma^2 \quad \bar{X} = \frac{\bar{Y} - \beta_1}{\beta_2} \quad \text{var}(\beta_2) = (SE(\beta_2))^2$$



Thank for your attention
Q&A