

- ÔN TẬP MÔN KINH TẾ LƯỢNG -

Hàm hồi quy tuyến tính (Phương pháp bình phương nhỏ nhất – OLS: Ordinary Least Squares):

PRF: $Y_i = \alpha + \beta X_i + u_i$.

SRF: $\hat{Y} = \hat{\alpha} + \hat{\beta} X_i$ (ước lượng)

- Tính giá trị trung bình mẫu (average value):

$$\bar{X} = \frac{\sum X_i}{n} \quad \text{và} \quad \bar{Y} = \frac{\sum Y_i}{n}$$

- Tính hệ số hồi quy (Coefficient):

$$\hat{\beta} = \frac{\sum X_i Y_i - n \cdot \bar{X} \cdot \bar{Y}}{\sum X_i^2 - n(\bar{X})^2} \quad \text{và} \quad \hat{\alpha} = \bar{Y} - \hat{\beta} \bar{X}$$

- Tính phương sai (Variance):

$$\sigma^2_Y = \frac{\sum (Y_i - \bar{Y})^2}{n-1} \quad \text{và} \quad \sigma^2_X = \frac{\sum (X_i - \bar{X})^2}{n-1}$$

- Tính độ lệch tiêu chuẩn (Standard Deviation):

$$SD_Y = \sqrt{\sigma^2_Y} \quad \text{và} \quad SD_X = \sqrt{\sigma^2_X}$$

- Tính khoảng phương sai hay hiệp phương sai (Covariance):

$$S_{XY} = \text{cov}(X, Y) = \frac{1}{n-1} * \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})$$

Tính tổng bình phương độ lệch:

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

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

$$RSS = \sum \hat{u}_i^2 = \sum (Y_i - \hat{Y}_i)^2$$

$$TSS = ESS + RSS$$

$$\text{Với } x_i = X_i - \bar{X} \quad \text{và} \quad y_i = Y_i - \bar{Y}$$

Tính hệ số xác định R^2 :

$$R^2 = 1 - \frac{RSS}{TSS} = \frac{ESS}{TSS} = \frac{\hat{\beta}^2 \sum x_i^2}{\sum y_i^2}$$

Với $0 < R^2 < 1$

$R^2=1$ hàm hồi quy thích hợp (mức độ hoàn hảo của mô hình) khi đó phần dư $RSS=0$

$$\Rightarrow \hat{Y}_i = Y_i, \forall i$$

$R^2=0 \Rightarrow$ SRF (mô hình hồi quy mẫu) không thích hợp $RSS=TSS \Rightarrow \hat{Y}_i = \bar{Y}, \forall i$

Hệ số tương quan: r (coefficient of Correlation)

$$r = \frac{\sum XiYi - n\bar{X}\bar{Y}}{\sqrt{\sum Xi^2 - n(\bar{X})^2} * \sqrt{\sum Yi^2 - n(\bar{Y})^2}}$$

Với $xi = Xi - \bar{X}$ và $yi = Yi - \bar{Y}$

Ta có thể viết:
$$r = \frac{\sum xi.yi}{\sqrt{\sum xi^2} \sqrt{\sum yi^2}} = \pm \sqrt{R^2}$$

r cùng dấu với $\hat{\beta}$

Tính khoảng tin cậy hệ số:

Bước 1: Xác định độ tin cậy 95% (hoặc 90%) để tìm được mức ý nghĩa $\alpha=5\%$ (hoặc 10%).

Tính $\alpha/2 = 0.025$. Tính giá trị t tra bảng t-student với phân vị $\alpha/2$ và bậc tự do $df=n-k-1$

Bước 2: Xác định phương sai PRF

$$\hat{\sigma}^2 = \frac{RSS}{n - k - 1}$$

Bước 3: Xác định sai số chuẩn (standard error) của từng hệ số.

$$s\hat{e}(\hat{\alpha}) = \sqrt{\frac{\sum Xi^2 * \hat{\sigma}^2}{n * \sum xi^2}} \quad \text{Với } xi = Xi - \bar{X}$$

$$s\hat{e}(\hat{\beta}) = \sqrt{\frac{\hat{\sigma}^2}{\sum xi^2}}$$

Bước 4: So sánh và tính khoảng tin cậy.

$$\hat{\alpha} \pm t_{\alpha/2}^{(n-k-1)} * s\hat{e}(\hat{\alpha}) \quad \text{hoặc} \quad \hat{\alpha} - t_{\alpha/2}^{(n-k-1)} * s\hat{e}(\hat{\alpha}) < \hat{\alpha} < \hat{\alpha} + t_{\alpha/2}^{(n-k-1)} * s\hat{e}(\hat{\alpha})$$

$$\hat{\beta} \pm t_{\alpha/2}^{(n-k-1)} * s\hat{e}(\hat{\beta}) \quad \text{hoặc} \quad \hat{\beta} - t_{\alpha/2}^{(n-k-1)} * s\hat{e}(\hat{\beta}) < \hat{\beta} < \hat{\beta} + t_{\alpha/2}^{(n-k-1)} * s\hat{e}(\hat{\beta})$$

Khoảng tin cậy của phương sai:

Bước 1: Xác định độ tin cậy 95% (hoặc 90%) để tìm được mức ý nghĩa $\alpha=5\%$ (hoặc 10%).

Tính phân vị $\alpha/2 = 0.025$ và $1-\alpha/2=0.975$. Tra bảng phân phối Chi-square với 2 phân vị $\alpha/2$ và

$1-\alpha/2$ cùng với bậc tự do $df=n-k-1$

$$X_{\alpha/2}^2(df) \text{ và } X_{1-\alpha/2}^2(df)$$

Bước 2: Tính khoảng tin cậy phương sai:
$$\sigma^2 = \left[\frac{(n-k-1)\hat{\sigma}^2}{X_{\alpha/2}^2(df)}; \frac{(n-k-1)\hat{\sigma}^2}{X_{1-\alpha/2}^2(df)} \right]$$

Kiểm định hệ số hồi quy:

- Bước 1: Đặt giả thuyết $H_0: \beta=0$ và đối thuyết $H_1: \beta \neq 0$ với mức ý nghĩa $\alpha=5\%$ (thông thường)
- Bước 2: Áp dụng 1 trong các cách sau:

Cách 1: Phương pháp khoảng tin cậy:

Kiểm định 2 phía: $[\hat{\theta} - t_{\alpha/2}^{(n-2)} * s\hat{e}(\hat{\theta}); \hat{\theta} + t_{\alpha/2}^{(n-2)} * s\hat{e}(\hat{\theta})]$

Nếu θ_0 không rơi vào khoảng này thì bác bỏ giả thuyết H_0 .

Kiểm định phía phải: $[\hat{\theta} - t_{\alpha/2}^{(n-2)} * s\hat{e}(\hat{\theta}); +\infty]$

Nếu θ_0 không rơi vào khoảng này thì bác bỏ giả thuyết H_0 .

Kiểm định phía trái: $[-\infty; \hat{\theta} + t_{\alpha/2}^{(n-2)} * s\hat{e}(\hat{\theta})]$

Nếu θ_0 không rơi vào khoảng này thì bác bỏ giả thuyết H_0 .

Cách 2: Phương pháp giá trị tới hạn:

Bước 1: Tính $t_0 = \frac{\hat{\beta} - \beta_0}{s\hat{e}(\hat{\beta})}$

Bước 2: Tra bảng với mức ý nghĩa $\alpha/2$ và α ($\alpha/2$ đối với kiểm định 2 phía và α đối với kiểm định 1 phía). Tra bảng t-student: $t_{\alpha/2}^{n-2}$ và t_{α}^{n-2}

Bước 3: So sánh t_0 với giá trị tới hạn.

Kiểm định 2 phía: $|t_0| > t_{\alpha/2}^{n-2}$: bác bỏ giả thuyết H_0 .

Kiểm định phía phải: $|t_0| > t_{\alpha}^{n-2}$: bác bỏ giả thuyết H_0 .

Kiểm định phía trái: $|t_0| < -t_{\alpha}^{n-2}$: bác bỏ giả thuyết H_0 .

Cách 3: Phương pháp giá trị P-value:

Bước 1: Tính giá trị $t_0 = \frac{\hat{\beta} - \beta_0}{s\hat{e}(\hat{\beta})}$

Bước 2: Tính P-value = $P(|t| > |t_0|)$

Bước 3: So sánh với mức ý nghĩa $\alpha=5\%$

Kiểm định 2 phía: $p\text{-value} < \alpha$: bác bỏ giả thuyết H_0 .

Kiểm định 1 phía: $p\text{-value}/2 < \alpha$: bác bỏ giả thuyết H_0 .

Kiểm định sự phù hợp của mô hình (F_0):

- R^2 càng gần 1, mô hình hồi quy càng có ý nghĩa. Do đó, cần đánh giá xem giá trị $R^2 > 0$ có ý nghĩa thống kê hay không.

- Nếu với mô hình hồi quy 2 biến, giả thuyết H_0 còn có ý nghĩa biến độc lập không ảnh hưởng đến biến phụ thuộc Y .

- Kiểm định bằng phương pháp giá trị tới hạn:

Bước 1: Đặt giả thuyết $H_0: R^2=0 \sim \beta=0$ và đối thuyết $H_1: R^2 > 0$

Bước 2: tính $F_0 = \frac{R^2(n-2)}{1-R^2} = \frac{ESS/1}{RSS/(n-2)}$

Bước 3: So sánh kết quả với $\alpha=5\%$. Tra bảng F với mức ý nghĩa α và 2 bậc tự do (1,n-2) ta tính giá trị tới hạn $F_{\alpha}(1,n-2)$.

So sánh F_0 và $F_{\alpha}(1,n-2)$

Nếu $F_0 > F_{\alpha}(1,n-2)$: bác bỏ giả thuyết H_0

Nếu $F_0 < F_{\alpha}(1,n-2)$: chấp nhận giả thuyết H_0 .

Đọc hiểu bảng kết quả hồi quy trên phần mềm Excel:

Regression

Statistics

Multiple R

Hệ số R có thể nhân đôi

R-Square (R²)

Hệ số xác định R²

$$R^2 = \frac{ESS}{TSS}$$

Ajusted R Square (r)

Hệ số tương quan r

$$r=1-[1-R^2]*(n-1/n-k-1)$$

Standard Error (σ)

Sai số chuẩn của PRF

$$\sqrt{\hat{\sigma}^2} = \sqrt{\frac{RSS}{n-k-df}}$$

Observation

Số quan sát

ANOVA

	df(bậc tự do)	SS (ESS)	MS(EMS)	F
Regression(ESS)		ESS	ESS/df (trung bình phần g.thích)	= $\frac{ESS / df}{RSS / df}$
Residual (RSS)		RSS	RSS/df (t.bình phần ko g.thích)	
Total (TSS)	TSS=ESS+RSS	TSS	TMS=EMS+RMS	

	Coefficient	standard error	t-stat	p-value	lower 95%	upper 95%
	Hệ số hồi quy	Sai số chuẩn (hồi quy)	t- thống kê	Giá trị P	Độ tin cậy (dưới)	Độ tin cậy (trên)
Intercept	$\hat{\alpha}$	$se(\hat{\alpha})$				
Variable 1 (biến 1)	$\hat{\beta}_2$	$se(\hat{\beta}_2)$	$t = \frac{\hat{\beta}_2 - \beta_0}{se(\hat{\beta}_2)}$			
Variable 1 (biến 2)	$\hat{\beta}_3$	$se(\hat{\beta}_3)$	$t = \frac{\hat{\beta}_3 - \beta_0}{se(\hat{\beta}_3)}$			

Đọc hiểu bảng kết quả hồi quy trên phần mềm Eviews:

Dependent Variable: CM Method: Least Squares Date: 08/18/07 Time: 21:46 Sample: 1 64 Included observations: 64				
	Số quan sát			
Variable	Coefficient	Std. Error	t-Statistic	Prob.
Biến trong mô hình	Hệ số HQ	Sai số chuẩn	Thống kê t	Giá trị P
C	$\hat{\alpha} = 263.6416$	$se(\hat{\alpha}) = 11.59318$		
PGNP	$\hat{\beta}_2 = -0.005647$	$se(\hat{\beta}_2) = 0.002003$	$t = \frac{\hat{\beta}_2 - \beta_0}{se(\hat{\beta}_2)}$	
R-squared (R^2)	hệ số xác định 0.707665	Mean dependent var ($\bar{Y}$)		141.5
Adjusted R-squared (R_{adj}) or $\bar{R}^2$	0.698081	S.D. dependent var $\sqrt{\frac{\sum (Y_i - \bar{Y})^2}{n-1}}$		75.97807
S.E. of regression ($\hat{\sigma}$) PRF	41.7478	Akaike info criterion (AIC)		10.34691
Sum squared resid (RSS)	106315.6	Schwarz criterion (SC)		10.44811
Log likelihood (L)	-328.1012	F-statistic	Giá trị thống kê F	73.83254
Durbin-Watson stat (DW)	2.186159	Prob(F-statistic) = P(phân phối F > Fo)		0.000000

Viết phương trình hồi quy:

Căn cứ vào kết quả hồi quy có trong bảng, ta có thể viết lại phương trình hồi quy mẫu như sau:

$$SRF: \hat{Y} = \hat{\alpha} + \hat{\beta}_2 X_i \quad (\text{ước lượng})$$

Trình bày kết quả hồi quy:

$$\hat{Y} = \hat{\alpha} + \hat{\beta}_2 X_i \quad n = ? \text{ (số quan sát?)}$$

$$t = \frac{\hat{\alpha} - \alpha_0}{se(\hat{\alpha})} \quad t = \frac{\hat{\beta}_2 - \beta_0}{se(\hat{\beta}_2)} \quad F_0 = ?$$

$$se(\hat{\alpha}) = ? \quad se(\hat{\beta}_2) = ? \quad R^2 = ?$$

$$p\text{-value(SRF)} = ? \quad P\text{-value (PRF)}$$

$$\hat{\sigma}^2 \text{ (PRF)} = ?$$

$$TSS = ? \quad ESS = ? \quad RSS = ?$$

Ý nghĩa hệ số hồi quy:

Đối với dạng hàm: $\hat{Y} = \hat{\alpha} + \hat{\beta}_2 X_i$ (hệ số hồi quy α, β có ý nghĩa là hệ số độ dốc)

Đối với dạng hàm $\log \hat{Y} = \hat{\alpha} + \hat{\beta}_2 \log X_i$ (hệ số hồi quy α, β có ý nghĩa là hệ số co giãn)

Đối với dạng hàm có biến giả: hệ số hồi quy β theo biến giả có ý nghĩa là hệ số cắt.

Ý nghĩa R^2 , F, DW.

$$R^2: R^2 = 1 - \frac{RSS}{TSS} = \frac{ESS}{TSS} = \frac{\hat{\beta}^2 \sum x_i^2}{\sum y_i^2} \quad (\text{Với } 0 < R^2 < 1)$$

$R^2 = 1$ dạng hoàn toàn thích hợp (mức độ hoàn hảo của mô hình) khi đó phần dư $RSS = 0 \Rightarrow$

$$\hat{Y}_i = Y_i, \forall i$$

$R^2 = 0 \Rightarrow$ SRF (mô hình hồi quy mẫu) không thích hợp $\rightarrow RSS = TSS \Rightarrow \hat{Y}_i = \bar{Y}_i, \forall i$

F: Giá trị thống kê F-stat = EMS/RMS (càng lớn càng tốt, chứng tỏ phần dư RSS nhỏ, mô hình phù hợp).

Durbin Waston stat (phương pháp OLS):

Sau khi xuất kết quả hồi quy, tìm phần dư e_i và tạo biến trễ phần dư e_{i-k} : độc lập.

$$DW = \frac{\sum (e_i - e_{i-k})^2}{\sum e_i^2} \quad \text{với } k=1$$

(Dùng để kiểm định mô hình có hay không có tương quan giữa các biến)

AIC: càng nhỏ càng tốt.

Quan hệ giữa R^2 và R^2_{adj} :

$$R^2 = 1 \Rightarrow R^2_{adj} = 1$$

$$R^2 = 0 \Rightarrow R^2_{adj} < 0 \text{ (R điều chỉnh có thể âm)}$$

Quan hệ giữa R^2 và F, R^2 và ESS, RSS.

$$F_0 = \frac{R^2(n-2)}{1-R^2} = \frac{ESS/1}{RSS/(n-2)}$$

Quan hệ giữa F và R^2 như sau: $F = \frac{ESS/k}{RSS/n-k-1} = \frac{R^2/k}{(1-R^2)/(n-k-1)}$ R^2 càng cao, F càng cao.

$$R^2 = 1 - \frac{RSS}{TSS} = \frac{ESS}{TSS} = \frac{\hat{\beta}^2 \sum x_i^2}{\sum y_i^2} \quad (\text{đo lường mức độ phù hợp của mô hình, dựa trên 2 biến}$$

chọn vào mô hình tuyến tính).

$$R^2_{adj} = 1 - \frac{RSS/(n-k)}{TSS/(n-1)} = 1 - \frac{(TSS - ESS)/(n-k)}{TSS/(n-1)} = 1 - (1-R^2) * \frac{n-1}{n-k}$$
 dùng cho các mô hình hồi quy có các biến giải thích khác nhau (xem mức độ thích hợp của biến).

Kiểm định giả thuyết đồng thời (kiểm định sự phù hợp của mô hình hồi quy đa biến):

Bước 1: Đặt giả thuyết $H_0: R^2=0 \sim H_0: \beta_1=\beta_2=0$ (ý nghĩa: các biến độc lập đồng thời không ảnh hưởng đến biến phụ thuộc hay nói cách khác: hàm hồi quy mẫu không phù hợp)
đối thuyết $H_1: R^2>0 \sim H_1: \text{có ít nhất một } \beta \neq 0$.

Bước 2: Tính giá trị F

$$F = \frac{ESS/(k-1)}{RSS/(n-k)} = \frac{R^2(n-k)}{(1-R^2)(k-1)} \sim F(k-1, n-k)$$

Bước 3: Tra bảng F với mức ý nghĩa $\alpha=5\%$ (thông thường) và phân vị $F(k-1, n-k)$.

Bước 4: So sánh kết quả giá trị F trong bảng kết quả hồi quy (F-statistic) với F tra bảng.

Kiểm định bằng phương pháp giá trị tới hạn: $F_0 > F_{\alpha}(k-1, n-k)$: bác bỏ giả thuyết H_0

Kiểm định bằng mức ý nghĩa α : $P\text{-value} = P(F > F_0) < \alpha$: bác bỏ giả thuyết H_0

Note: F_0 càng cao thì khả năng bác bỏ giả thuyết H_0 càng lớn.

Kiểm định Wald Test.

Ý nghĩa: xem xét có nên đưa thêm biến mới vào mô hình hay không?

Xét 2 mô hình:

Mô hình ràng buộc (UR-unrestricted model): $Y = \beta_0 + \beta_1 X_1 + \dots + \beta_{m-1} X_{m-1} + \dots + \beta_{k-1} X_{k-1} + u_i$.

Mô hình ràng buộc (R – restricted model) : $Y = \beta_0 + \beta_1 X_1 + \dots + \beta_{m-1} X_{m-1} + u_i$.

Kiểm định bằng thống kê F:

Bước 1: ước lượng mô hình UR với k tham số, lấy kết quả của RSS có $df=n-k$

Ước lượng mô hình R với m tham số, lấy kết quả của RSS có $df=n-m$.

Trong đó: m là số ràng buộc, $m=k_1-k_2$

k_2 là số biến giải thích trong mô hình R

k_1 là số biến giải thích trong mô hình UR

Bước 2: Tra bảng F với mức ý nghĩa $\alpha=5\%$ (thông thường) và $F_{\alpha}(k-m, n-k)$.

$$\text{Tính } F_{tt} = \frac{(RSS_R - RSS_{UR})/(k-m)}{RSS_{UR}/(n-k)} = \frac{(R_{UR}^2 - R_R^2)/(k-m)}{(1 - R_{UR}^2)/(n-k)}$$

Bước 3: So sánh F tính toán với F tra bảng.

$F_{tt} > F_{\alpha}(k-m, n-k)$: bác bỏ giả thuyết H_0 (nên đưa biến vào mô hình)

$F_{tt} < F_{\alpha}(k-m, n-k)$: chấp nhận giả thuyết H_0 (không nên đưa biến vào mô hình)

Kiểm định Chow Test:

Ý nghĩa: Xem trong chuỗi dữ liệu có khác nhau gì về cấu trúc không?

Nếu khác tách thành các mô hình khác nhau.

Nếu giống chỉ dùng một mô hình.

Ý tưởng: có nên tách riêng hay để chung mô hình

Thực hiện:

Bước 1: ước lượng 3 mô hình

$Y = \alpha_1 + \alpha_2 X + v_1$. trong giai đoạn đầu có n_1 quan sát (VD: 1997~1990)

Tính RSS_1 với $df = n_1 - k$

$Y = \beta_1 + \beta_2 X + v_2$. trong giai đoạn sau có n_2 quan sát (VD: 1991~1998)

Tính RSS_2 với $df = n_2 - k$ (k là tham số của mô hình hồi quy)

Đặt $RSS_U = RSS_1 + RSS_2$ với bậc tự do $df = n_1 + n_2 - 2k$

Ước lượng mô hình chung $Y = \gamma_1 + \gamma_2 X + u$ với số quan sát $n = n_1 + n_2$

Tính RSS_R với $df = n - k$

Bước 2: Tính giá trị của F-statistic $F_{tt} = \frac{(RSS_R - RSS_{UR})/k}{RSS_{UR}/(n-2k)}$

Bước 3: kiểm định

Giả thuyết H_0 : hai hồi quy của 2 thời kỳ như nhau

Đối thuyết H_1 : hai hồi quy khác nhau.

$F_{tt} > F_{\alpha}(k, n-2k)$: bác bỏ giả thuyết H_0

$F_{tt} < F_{\alpha}(k, n-2k)$: chấp nhận giả thuyết H_0

Xác định biến giả:

Cách tạo biến giả:

Đối với dữ liệu chéo, biến giả có thể theo giai đoạn

D=0: giai đoạn 1

D=1: giai đoạn 2

Bảng Eviews:

Cách 1: nhập giá trị 0,1 vào các quan sát tương ứng.

Cách 2: * tạo biến xu thế Eviews/genr/tt=@trend(mốc cuối giai đoạn 1)

* tạo biến giả dựa trên biến xu thế, Eviews/genr/DUM=tt>số quan sát.

Đối với 2 thuộc tính: D=1 (thuộc tính trái), phần còn lại D=0 (biến không có trong mô hình)

Đối với nhiều thuộc tính, số biến giả = số thuộc tính -1. So sánh các thuộc tính khác với thuộc tính cơ sở. Tính % khác biệt của biến giả bằng cách lấy 1-antilog

Kiểm định:

Phương pháp khoảng tin cậy (liên hệ phần tính khoảng tin cậy)

Phương pháp mức ý nghĩa: (liên hệ kiểm định bằng giá trị P-value với mức ý nghĩa)

Phương pháp nên hay không đưa biến vào mô hình (kiểm định bằng thống kê F)

Note: Ta cần chú ý đến mô hình hồi quy trước và sau khi có biến giả để đánh giá. Khi đưa biến giả vào mô hình, các hệ số hồi quy có ý nghĩa (R^2 , t-stat và P-value) sẽ cho ta nhận định đúng hơn về mô hình. Khi đó mới kết luận mô hình phù hợp hay không.

Phát hiện phương sai thay đổi:Phát hiện:

Để phát hiện phương sai của nhiễu có thay đổi hay không, người ta thường dùng công cụ chuẩn đoán phần dư U_i (có thể có kết quả đáng tin cậy).

Trong dữ liệu chéo đó lấy mẫu rất rộng, để suy ra phương sai thay đổi.

Phân tích phần dư U_i , và vẽ đồ thị phần dư theo biến độc lập bất kỳ, ta có dạng hình phân tán đều và đồng nhất.

Kiểm định Park test

Bước 1: Hồi quy mô hình, lấy số liệu phần dư (resid trong bảng biến tại phần mềm Eviews).

Mô hình (1): $Y_i = \beta_1 + \beta_2 X_i + U_i$

Bước 2: ước lượng mô hình phần dư theo biến độc lập.

Mô hình (2): $\ln U_i = \alpha_1 + \alpha_2 X_i + V_i$.

Bước 3: đặt giả thuyết $H_0: \alpha_2 = 0$ (phương sai ko đổi)

Đối thuyết $H_1: \alpha_2 \neq 0$ (phương sai thay đổi)

Kiểm định bằng t-stat.

Kiểm định Glejsei test

Bước 1: hồi quy mô hình, lấy số liệu phần dư (resid trong bảng biến tại phần mềm Eviews).

Mô hình (1): $Y_i = \beta_1 + \beta_2 X_i + U_i$

Bước 2: ước lượng mô hình phần dư theo biến độc lập.

Mô hình (2) có 1 trong các dạng sau :

$$|\hat{U}_i| = \alpha_1 + \alpha_2 X_i + V_i \quad \text{hoặc} \quad |\hat{U}_i| = \alpha_1 + \alpha_2 \frac{1}{X_i} + V_i$$

$$|\hat{U}_i| = \alpha_1 + \alpha_2 \frac{1}{\sqrt{X_i}} + V_i \quad \text{hoặc} \quad |\hat{U}_i| = \alpha_1 + \alpha_2 \sqrt{X_i} + V_i$$

Bước 3: đặt giả thuyết $H_0: \alpha_2 = 0$ (phương sai không đổi)

Đối thuyết $H_1: \alpha_2 \neq 0$ (phương sai thay đổi)

Kiểm định bằng t-stat.

Kiểm định White test:

Bước 1: hồi quy mô hình, lấy số liệu phần dư (resid trong bảng biến tại phần mềm Eviews).

Mô hình (1): $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + U_i$

Bước 2: ước lượng mô hình phải bằng thao tác Eviews (View/Residual Tests/White Heteroscedasticity) thu được R^2 . Sau đó ta tính $X_{tt} = n \cdot R^2$ (trong đó n là số quan sát)

Bước 3: đặt giả thuyết $H_0: \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$ (phương sai không đổi)

Đối thuyết $H_1: \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 \neq 0$ (phương sai thay đổi)

Bước 4: kiểm định và so sánh.

Tra bảng Chi-square $X^2_\alpha(df)$ với mức ý nghĩa α

Nếu $X_{tt} = n \cdot R^2 > X_{tt} = n \cdot R^2$: bác bỏ giả thuyết.

Phát hiện tự tương quan bằng kiểm định Durbin Waston:

Phát hiện: căn cứ vào đồ thị Scatter của phần dư U_i với biến trễ U_{i-1} .

- Đồ thị có dạng ngẫu nhiên thì không có tự tương quan.
- Đồ thị có dạng hệ thống thì nhận định có tự tương quan xảy ra.

Thực hiện kiểm định bằng Durbin Waston

Bước 1: ước lượng mô hình hồi quy gốc. lấy giá trị phần dư U_i và tạo biến trễ U_{i-1} .

Bước 2: Tính giá trị
$$\rho = \frac{\sum_{i=2}^n \hat{U}_i \hat{U}_{i-1}}{\sum_{i=1}^n \hat{U}_i^2} \quad \text{với } -1 \leq \rho \leq 1$$

Hoặc tính giá trị
$$d = \frac{\sum_{i=2}^n (\hat{U}_i - \hat{U}_{i-1})^2}{\sum_{i=1}^n \hat{U}_i^2} \approx 2(1 - \hat{\rho}) \quad \text{với } 0 \leq d \leq 4$$

Bước 3: kiểm định và so sánh

Tra bảng thống kê Durbin Waston cho ta các giá trị tới hạn d_U và d_L với mức ý nghĩa α , số quan sát n và số biến độc lập k .

So sánh:

- * $d \in (0, d_L)$: tự tương quan dương (thuận chiều)
- * $d \in (d_L, d_U)$: không quyết định được
- * $d \in (d_U, 2)$: không có tương quan bậc nhất.
- * $d \in (2, 4 - d_U)$: không có tương quan bậc nhất.
- * $d \in (4 - d_U, 4 - d_L)$: không quyết định được
- * $d \in (4 - d_L, 4)$: tự tương quan âm.

Phát hiện đa cộng tuyến:

Phát hiện: R^2 cao nhưng t-stat thấp (không có ý nghĩa P-value có giá trị cao)

Hệ số tương quan cặp giữa các biến giải thích cao, khoảng 0.8

Thực hiện kiểm định và xác định đa cộng tuyến:

Bước 1: xét hệ số tương quan giữa 2 biến (có đa cộng tuyến)

Nếu hệ số tương quan gần bằng 1 (đa cộng tuyến gần như hoàn hảo)

Nếu hệ số tương quan < 0.8 (đa cộng tuyến không hoàn hảo).

Bước 2: Hồi quy Y theo từng biến độc lập X_1, X_2 .

Ta có 2 mô hình:

(1): $Y^*_1 = \alpha + \alpha_1 X_1$ lấy kết quả R^2 , p-value (xem có hay không ý nghĩa thống kê)

(2): $Y^*_2 = \beta + \beta_2 X_2$ lấy kết quả R^2 , p-value (xem có hay không ý nghĩa thống kê)

Bước 3: Hồi quy mô hình phải 2 biến có đa cộng tuyến

(3) $X^*_2 = \gamma + \gamma_1 X_1$ lấy kết quả R^2 , p-value (xem có hay không ý nghĩa thống kê)

Bước 4: đặt giả thuyết H_0 : không có đa cộng tuyến

Đối thuyết H_1 : có đa cộng tuyến

Kiểm định bằng thống kê F:

$$F_2 = \frac{R_2^2 / (k - 2)}{(1 - R_2^2) / (n - k + 1)}$$

Tính F tra bảng với mức ý nghĩa α , $F_{\alpha}(k-2, n-k+1)$

So sánh: $F_2 > F_{\alpha}(k-2, n-k+1)$: bác bỏ giả thuyết.

$F_2 < F_{\alpha}(k-2, n-k+1)$: chấp nhận giả thuyết.

Thực hiện kiểm định và bỏ bớt biến:

Bước 1: xét hệ số tương quan giữa 2 biến (có đa cộng tuyến)

Nếu hệ số tương quan gần bằng 1 (đa cộng tuyến gần như hoàn hảo)

Nếu hệ số tương quan < 0.8 (đa cộng tuyến không hoàn hảo).

Bước 2: Hồi quy Y theo từng biến độc lập X_1, X_2 .

Ta có 2 mô hình:

(1): $Y^*_1 = \alpha + \alpha_1 X_1$ lấy kết quả R^2 , p-value (xem có hay không ý nghĩa thống kê)

(2): $Y^*_2 = \beta + \beta_2 X_2$ lấy kết quả R^2 , p-value (xem có hay không ý nghĩa thống kê)

Bước 3: kiểm định

Xét p-value của X_1 và p-value của X_2 trong kết quả hồi quy.

p-value (X_1) > p-value (X_2): mô hình hồi quy Y theo X_1 có mức độ phù hợp cao hơn mô hình hồi quy Y theo X_2 . Do đó loại bỏ biến X_2 .

Cách khắc phục các loại bệnh (phương sai thay đổi, tự tương quan, đa cộng tuyến):

Cách khắc phục đa cộng tuyến:

Bỏ biến ra khỏi mô hình, sau hồi quy lại mô hình không bao gồm biến cần loại bỏ. Đánh giá giá trị R^2 , t-stat và P-value xem có ý nghĩa thống kê không.

Căn cứ vào kết quả earnings (hệ số đáng tin cậy cho trước). Sau đó xác định mô hình hồi quy phụ theo hệ số cho trước. Đánh giá giá trị R^2 , t-stat và P-value của mô hình hồi quy phụ xem

có ý nghĩa thống kê không.

Thêm dữ liệu cho mô hình, tuy nhiên cách thức này toán kèm chi phí nên ít được thực hiện.

Cách khắc phục phương sai thay đổi:

- Biết phương sai σ^2
- Không biết phương sai σ^2 :

Bước 1: ước lượng phương trình (1): $Y_i = b_1 + b_2 X_i + u_i$

Bước 2: vẽ đồ thị phần dư u_i theo X_i . Đánh giá xem phương sai nhiễu có hay không tỷ lệ thuận với biến giải thích.

Bước 3: Chia 2 vế của phương trình hồi quy (1) cho căn bậc 2 của biến giải thích.

$$(2) \frac{Y_i}{\sqrt{X_i}} = \frac{b_1}{\sqrt{X_i}} + b_2 \frac{X_i}{\sqrt{X_i}} + \frac{u_i}{\sqrt{X_i}} \Leftrightarrow \frac{Y_i}{\sqrt{X_i}} = \frac{b_1}{\sqrt{X_i}} + b_2 \sqrt{X_i} + v_i$$

Chuyển thành dạng phương trình ko có hệ số cắt.

Bước 4: So sánh mô hình (1) và (2) qua số liệu hồi quy R^2 , t-stat và P-value và đánh giá mô hình.

Cách khắc phục tự tương quan:

- Trường hợp biết cấu trúc của tự tương quan.
- Trường hợp chưa biết cấu trúc của tự tương quan:

Cách 1: ước lượng ρ bằng thống kê d

Cách 2: phương pháp Durbin Waston 2 bước (sách KTL-trang 171)