KINH TẾ LƯỢNG CƠ BẢN – BASIC ECONOMETRICS

Bài Mở Đầu

1. Khái niệm về Kinh tế lượng (Econometrics)

- *Econo + Metric*

Khái niệm: KTL nghiên cứu những mối quan hệ Kinh tế Xã hội; thông qua việc xây dựng, phân tích, đánh giá các mô hình để cho ra lời giải bằng số, hỗ trợ việc ra quyết định.

- KTL sử dụng kết quả của :
  + Lý thuyết kinh tế
  + Mô hình toán kinh tế
  + Thống kê, xác suất

2. Phương pháp luận

2.1. Đặt giả thiết về vấn đề nghiên cứu

- Xác định phạm vi, bản chất, tính chất của các đối tượng và mối quan hệ giữa chúng.

2.2. Xây dựng mô hình phù hợp

- Xác định mô hình lý thuyết kinh tế hợp lý.
- Xây dựng mô hình toán kinh tế :
  + Mỗi đối tượng đại diện bởi một hoặc một số biến số.
  + Mỗi mối quan hệ: Phương trình, hàm số, bất phương trình…
  + Giá trị các tham số : cho biết bản chất mối quan hệ.

2.3. Thu thập số liệu và ước lượng tham số

- Số liệu được dùng : từ thống kê.
- Bằng phương pháp cụ thể : ước lượng các tham số.

Với bộ số liệu xác định và phương pháp cụ thể, kết quả ước lượng là những con số cụ thể.

2.4. Kiểm định

- Bằng phương pháp kiểm định thống kê: kiểm định giá trị các tham số, bản chất mối quan hệ
- Kiểm định tính chính xác của mô hình.
- Nếu không phù hợp : quay lại các bước trên.
- Biến đổi, xây dựng mô hình mới để có kết quả tốt nhất.

2.5. Dự báo

- Dựa trên kết quả được cho là tốt : dự báo về mối quan hệ, về các đối tượng trong những điều kiện xác định.
- Đánh giá quyết định.

3. Số liệu dùng trong KTL

3.1. Phân loại

- Số liệu theo thời gian.
- Số liệu theo không gian.
- Số liệu chéo

3.2. Nguồn gốc

- Điều tra
- Mua
- Từ nguồn được phát hành : Niên giám thống kê

3.3. Tính chất của số liệu

- Số liệu ngẫu nhiên phi thực nghiệm.
- Phù hợp mục đích nghiên cứu.

Chương 1. CÁC KHÁI NIỆM CƠ BẢN

1. Phân tích hồi qui (*Regression*)

1.1. Định nghĩa

*Phân tích hồi qui là phân tích mối liên hệ phụ thuộc giữa **một biến** gọi **là biến phụ thuộc** (biến được giải thích, biến nội sinh) phụ thuộc vào **một hoặc một số biến** khác gọi là (các) **biến giải thích** (biến độc lập, biến ngoại sinh, biến hồi qui).*

1.2. Ví dụ

- Biến phụ thuộc (*dependent variable*) : Y
- Biến giải thích / hồi qui (*regressor(s)*) : X, hoặc X_2, X_3….
- Biến giải thích nhận những giá trị xác định, trong điều kiện đó biến phụ thuộc là một đại lượng ngẫu nhiên (biến ngẫu nhiên).

Phân tích hồi qui nghiên cứu mối liên hệ phụ thuộc giữa đại lượng ngẫu nhiên biến phụ thuộc phụ thuộc vào các giá trị xác định của (các) biến giải thích như thế nào.

$$X = X_i \rightarrow (Y/X_i)$$

1.3. Mục đích hồi qui

- **Ước lượng** (*Estimate*) trung bình biến phụ thuộc và các tham số.
- **Kiểm định** (*Hypothesis testing*) về mối quan hệ.
- **Dự báo** (*Forecast, Prediction*) giá trị biến phụ thuộc khi biến giải thích thay đổi.

(*)*Hồi qui : qui về trung bình*

1.4. So sánh với các quan hệ toán khác

- Quan hệ hàm số : $x \rightarrow !\, y$
- Quan hệ qua hệ số tương quan ρ_{xy}
- Quan hệ nhân quả $X \leftrightarrow Y$

2. Mô hình hồi qui Tổng thể

- Phân tích hồi qui dựa trên toàn bộ tổng thể
- Để thuận tiện trong phần này: biến phụ thuộc Y phụ thuộc một biến giải thích X

2.1. Hàm hồi qui tổng thể (PRF : *Population Regression Function*).

$X = X_i \rightarrow (Y/X_i)$

$\rightarrow \exists\, F(Y/X_i)$

$\rightarrow \exists\, !\, E(Y/X_i)$

$X_i \rightarrow !\, E(Y/X_i)$

$E(Y/X_i) = f(X_i)$ hoặc $E(Y/X) =$ Hàm hồi qui tổng thể (PRF)

$f(X)$

Nếu: hàm hồi qui tổng thể có dạng

$$\boldsymbol{E(Y/X) = \beta_1 + \beta_2 X}$$

Thì $\beta_1 = E(Y/X = 0)$: **hệ số chặn** (INPT : *intercept term*)

$\beta_2 = \dfrac{\partial E(Y/X)}{\partial X}$: **hệ số góc** (*slope coefficient*)

→ PRF cho biết quan hệ giữa biến phụ thuộc và biến giải thích về mặt trung bình trong tổng thể.

2.2. Phân loại

Hàm hồi qui tổng thể được gọi là **tuyến tính** nếu nó tuyến tính với tham số.

2.3. Yếu tố ngẫu nhiên

- Giá trị cụ thể $Y_i \in (Y/X_i)$, thông thường $Y_i \neq E(Y/X_i)$
- Đặt $u_i = Y_i - E(Y/X_i)$: là **yếu tố ngẫu nhiên** (nhiễu, sai số ngẫu nhiên: *random errors*)
- Tính chất của YTNN : + Nhận những giá trị dương và âm.
  + Kì vọng bằng 0: $E(u_i) = 0 \;\; \forall i$

Bản chất của YTNN : đại diện cho tất cả những yếu tố không phải biến giải thích nhưng cũng tác động tới biến phụ thuộc:

+ Những yếu tố không biết.
+ Những yếu tố không có số liệu.
+ Những yếu tố mà tác động của nó quá nhỏ không mang tính hệ thống.

3. Mô hình hồi qui mẫu

- Không biết toàn bộ Tổng thể, nên dạng của PRF có thể biết nhưng giá trị β_j thì **không biết**.
- Mẫu : một bộ phận mang thông tin của tổng thể.
- $W = \{(X_i, Y_i), i = 1 \div n\}$ được gọi là một **mẫu kích thước *n*, *n* quan sát** (*observation*).

3.1. Hàm hồi qui mẫu (SRF : *Sample Regression Function*)

Trong mẫu W, tồn tại một hàm số mô tả xu thế biến động của biến phụ thuộc theo biến giải thích về mặt trung bình, $\hat{Y} = \hat{f}(X)$ gọi là hàm hồi qui mẫu (SRF).

Hàm hồi qui mẫu có dạng giống hàm hồi qui tổng thể

Nếu PRF có dạng $E(Y/X_i) = \beta_1 + \beta_2 X_i$

Thì SRF có dạng $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$

- Vì có vô số mẫu ngẫu nhiên, nên có vô số giá trị của $\hat{\beta}_1$ và $\hat{\beta}_2 \rightarrow \hat{\beta}_j$ là biến ngẫu nhiên.
- Với một mẫu cụ thể w kích thước n, $\hat{\beta}_j$ sẽ là con số cụ thể.

3.2. Phần dư

Thông thường $Y_i \neq \hat{Y}_i$, đặt $e_i = Y_i - \hat{Y}_i$ và gọi là **phần dư** (*residual*).

Bản chất của phần dư e_i giống yếu tố ngẫu nhiên u_i

$\hat{Y}_i$, $\hat{\beta}_1, \hat{\beta}_2$, e_i là ước lượng điểm tương ứng của $E(Y/X_i)$, β_1, β_2, u_i.

Tóm tắt chương

$$E(Y/X_i) = \beta_1 + \beta_2 X_i$$

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

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

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

Trường hợp tổng quát

$$E(Y_i) = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \ldots + \beta_k X_{ki}$$

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \ldots + \beta_k X_{ki} + u_i$$

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \ldots + \hat{\beta}_k X_{ki}$$

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \ldots + \hat{\beta}_k X_{ki} + e_i$$

Chương 2. ƯỚC LƯỢNG VÀ PHÂN TÍCH MÔ HÌNH HỒI QUI HAI BIẾN

1. Mô hình

- Mô hình có dạng: $E(Y/X_i) = \beta_1 + \beta_2 X_i$

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

- Với **mẫu kích thước *n*** : $W = \{(X_i, Y_i), i = 1 \div n\}$, tìm $\hat{\beta}_1, \hat{\beta}_2$ sao cho SRF: $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$ phản ánh xu thế biến động về mặt trung bình của mẫu.

2. Phương pháp bình phương nhỏ nhất (OLS – *Ordinary Least Square*)

2.1.Phương pháp

- Tìm $\hat{\beta}_1$, $\hat{\beta}_2$ sao cho $\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 = \sum_{i=1}^{n} e_i^2 \to \min$

$$\hat{\beta}_2 = \frac{\overline{XY} - \overline{X}\overline{Y}}{\overline{X^2} - (\overline{X})^2} \qquad \hat{\beta}_1 = \overline{Y} - \hat{\beta}_2 \overline{X}$$

Đặt $x_i = X_i - \overline{X}$; $y_i = Y_i - \overline{Y}$ $\Rightarrow \hat{\beta}_2 = \dfrac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2 y_i}$

$\hat{\beta}_1$, $\hat{\beta}_2$ ước lượng bằng phương pháp bình phương nhỏ nhất, gọi là các ước lượng bình phương nhỏ nhất (OLS) của β_1 và β_2.

- Một số tính chất của các ước lượng bình phương nhỏ nhất:

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

2.2.Các giả thiết OLS

Để ước lượng OLS là tốt nhất thì tổng thể phải thỏa mãn một số giả thiết sau

Giả thiết 1: Biến giải thích là phi ngẫu nhiên

Giả thiết 2: Trung bình yếu tố ngẫu nhiên bằng 0 $E(u_i) = 0 \ \forall i$

Giả thiết 3: Phương sai yếu tố ngẫu nhiên bằng nhau $Var(u_i) = Var(u_j) = \sigma^2 \ \forall i \neq j$

Giả thiết 4: Các yếu tố ngẫu nhiên không tương quan $Cov(u_i, u_j) = 0 \ \forall i \neq j$

Giả thiết 5: YTNN và biến giải thích không tương quan $Cov(u_i, X_i) = 0 \ \forall i$

Định lý Gauss-Markov: *Nếu mô hình hồi quy thỏa mãn các giả thiết trên thì ước lượng OLS sẽ là ước lượng* ***tuyến tính, không chệch, tốt nhất*** *(trong số các ước lượng không chệch) của các tham số.*

2.3.Các tham số của ước lượng OLS

Các ước lượng $\hat{\beta}_j$ là biến ngẫu nhiên tùy thuộc mẫu, nên có các tham số đặc trưng

Kì vọng : $E(\hat{\beta}_1) = \beta_1$ $\qquad E(\hat{\beta}_2) = \beta_2$

Phương sai : $Var(\hat{\beta}_1) = \dfrac{\sum_{i=1}^{n} X_i^2}{n\sum_{i=1}^{n} x_i^2}\sigma^2$; $\quad Var(\hat{\beta}_2) = \dfrac{1}{\sum_{i=1}^{n} x_i^2}\sigma^2$

Độ lệch chuẩn : $Se(\hat{\beta}_j) = \sqrt{Var(\hat{\beta}_j)}$ $(j = 1,2)$

Hiệp phương sai: $Cov(\hat{\beta}_1, \hat{\beta}_2) = -\overline{X}Var(\hat{\beta}_2)$.

Với σ^2 là phương sai yếu tố ngẫu nhiên chưa biết, ước lượng bởi $\hat{\sigma}^2$: $\hat{\sigma}^2 = \dfrac{\sum_{i=1}^{n} e_i^2}{n-k}$

với ***k* là số tham số cần phải ước lượng** của mô hình.

$\hat{\sigma} = \sqrt{\hat{\sigma}^2}$ là **độ lệch chuẩn của đường hồi qui** : (*Se. of Regression*)

3. Phân tích các hệ số

Giả thiết: YTNN có phân phối chuẩn : $u_i \sim N(0; \sigma^2)$ $\forall i$, *khi đó:*

$$\hat{\beta}_j \sim N(\beta_j; Var(\hat{\beta}_j)); \qquad \frac{(n-k)\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-k); \qquad Y_i \sim N(\beta_1 + \beta_2 X_i, \sigma^2).$$

3.1. Ước lượng khoảng

Với độ tin cậy $(1 - \alpha)$ cho trước, ta có

i. Khoảng tin cậy cho các hệ số hồi quy

KTC đối xứng : $\hat{\beta}_j - Se(\hat{\beta}_j)t_{\alpha/2}(n-k) < \beta_j < \hat{\beta}_j + Se(\hat{\beta}_j)t_{\alpha/2}(n-k)$

KTC tối đa: $\beta_j < \hat{\beta}_j + Se(\hat{\beta}_j)t_{\alpha}(n-k)$

KTC tối thiểu: $\hat{\beta}_j - Se(\hat{\beta}_j)t_{\alpha}(n-k) < \beta_j$

ii.Khoảng tin cậy cho phương sai yếu tố ngẫu nhiên

KTC 2 phía: $\dfrac{\hat{\sigma}^2(n-k)}{\chi^2_{\alpha/2}(n-k)} < \sigma^2 < \dfrac{\hat{\sigma}^2(n-k)}{\chi^2_{1-\alpha/2}(n-k)}$

KTC tối đa $\sigma^2 < \dfrac{\hat{\sigma}^2(n-k)}{\chi^2_{1-\alpha}(n-k)}$

KTC tối thiểu $\dfrac{\hat{\sigma}^2(n-k)}{\chi^2_{\alpha}(n-k)} < \sigma^2$

3.2. Kiểm định giả thiết

Với mức ý nghĩa α cho trước

i. Kiểm định giả thiết cho các hệ số hồi quy

Cặp giả thiết	Tiêu chuẩn kiểm định	Miền bác bỏ H_0
$\begin{cases} H_0: \beta_j = \beta_j^* \\ H_1: \beta_j \neq \beta_j^* \end{cases}$	$T_{qs} = \dfrac{\hat{\beta}_j - \beta_j^*}{Se(\hat{\beta}_j)}$	$\lvert T_{qs} \rvert > t_{\alpha/2}(n-k)$
$\begin{cases} H_0: \beta_j = \beta_j^* \\ H_1: \beta_j > \beta_j^* \end{cases}$		$T_{qs} > t_{\alpha}(n-k)$

$\begin{cases} H_0 : \beta_j = \beta_j^* \\ H_1 : \beta_j < \beta_j^* \end{cases}$		$T_{qs} < -\ t_{\alpha}(n-k)$

Trường hợp đặc biệt $\begin{cases} H_0 : \beta_2 = 0 \\ H_1 : \beta_2 \neq 0 \end{cases}$ $\rightarrow T_{qs} = \dfrac{\hat{\beta}_2}{Se(\hat{\beta}_j)}$

*** Dùng** $P-value$

$H_1 : \beta_j < \beta_j^* \quad \Rightarrow P-value = P(t < t_{qs})$

$H_1 : \beta_j > \beta_j^* \quad \Rightarrow P-value = P(t > t_{qs})$

$H_1 : \beta_j \neq \beta_j^* \quad \Rightarrow P-value = 2P(t > | t_{qs} |)$

Nếu $P-value < \alpha$ thì bác bỏ giả thiết H_0

Nếu $P-value > \alpha$ thì chưa có cơ sở bác bỏ giả thiết H_0.

ii.Kiểm định giả thiết cho phương sai yếu tố ngẫu nhiên

Cặp giả thiết	Tiêu chuẩn kiểm định	Miền bác bỏ H_0
$\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 \neq \sigma_0^2 \end{cases}$	$\chi_{qs}^2 = \dfrac{(n-k)\hat{\sigma}^2}{\sigma_0^2}$	$\left[\begin{array}{l} \chi_{qs}^2 < \chi_{1-\alpha/2}^2(n-k) \\ \chi_{qs}^2 > \chi_{\alpha/2}^2(n-k) \end{array}\right.$
$\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 > \sigma_0^2 \end{cases}$		$\chi_{qs}^2 > \chi_{\alpha}^2(n-k)$
$H_0 : \sigma^2 = \sigma_0^2$ $H_1 : \sigma^2 < \sigma_0^2$		$\chi_{qs}^2 < \chi_{1-\alpha}^2(n-k)$

Chú ý

+) Giả thiết H_0 bao giờ cũng chứa dấu "=".

+) Chú ý khi tìm khoảng tin cậy và xây dựng cặp giả thiết với các hệ số β_j âm.

4. Sự phù hợp của hàm hồi qui

4.1. Hệ số xác định R^2

$$\left.\begin{array}{l} y_i = Y_i - \bar{Y} \\ \hat{y}_i = \hat{Y}_i - \bar{Y} \\ e_i = Y_i - \hat{Y} \end{array}\right\} \quad y_i = \hat{y}_i + e_i \ ; \ \text{Và chứng minh được} \ \sum_{i=1}^{n} y_i^2 = \sum_{i=1}^{n} \hat{y}_i^2 + \sum_{i=1}^{n} e_i^2$$

$$TSS = ESS + RSS$$

TSS (*Total Sum of Squares*) : đo tổng biến động của biến phụ thuộc

ESS (*Explained Sum of Squares*): tổng biến động của biến phụ thuộc được giải thích bởi mô hình (biến giải thích.).

RSS (*Residual SS*) : tổng biến động của biến phụ thuộc được giải thích bởi các yếu tố nằm ngoài mô hình – Yếu tố ngẫu nhiên.

Đặt $\boldsymbol{R^2} = \dfrac{ESS}{TSS} = 1 - \dfrac{RSS}{TSS}$ gọi là **hệ số xác định**, $0 \leq R^2 \leq 1$

***Ý nghĩa**: Hệ số xác định R^2 là tỉ lệ (hoặc tỉ lệ %) sự biến động của biến phụ thuộc được giải thích bởi biến giải thích (theo mô hình, trong mẫu).*

4.2. Kiểm định sự phù hợp của hàm hồi quy

Cặp giả thiết

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

Kiểm định F: $$F_{qs} = \frac{ESS/(k-1)}{RSS/(n-k)} = \frac{R^2}{1-R^2} \times \frac{n-k}{k-1}$$

- Nếu $F_{qs} > F_{\alpha}(k - 1; n - k)$ thì bác bỏ H_0 : hàm hồi qui được gọi là phù hợp.
- Ngược lại, hàm hồi qui không phù hợp.

Chú ý: Với mô hình hồi quy đơn ($k = 2$), ta có

- $\begin{cases} H_0 : R^2 = 0 \\ H_1 : R^2 > 0 \end{cases} \Leftrightarrow \begin{cases} H_0 : \beta_2 = 0 \\ H_1 : \beta_2 \neq 0 \end{cases}$
- ***Giá trị: $F_{qs} = \left(\frac{\hat{\beta}_2}{Se(\hat{\beta}_j)}\right)^2$.***

5. Dự báo

Là ước lượng khoảng cho giá trị trung bình và cá biệt của biến phụ thuộc khi biến giải thích nhận giá trị xác định $X = X_0$

a. Dự báo giá trị trung bình

$$\hat{Y}_0 - Se(\hat{Y}_0)t_{\alpha/2}(n-k) < E(Y/X_0) < \hat{Y}_0 + Se(\hat{Y}_0)t_{\alpha/2}(n-k)$$

Với $\hat{Y}_0 = \hat{\beta}_1 + \hat{\beta}_2 X_0$ và $Se(\hat{Y}_0) = \hat{\sigma}\sqrt{\frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\Sigma x_i^2}}$

b. Dự báo giá trị cá biệt

$$\hat{Y}_0 - Se(Y_0)t_{\alpha/2}(n-k) < Y_0 < \hat{Y}_0 + Se(Y_0)\, t_{\alpha/2}(n-k)$$

Với $Se(Y_0) = \hat{\sigma}\sqrt{1 + \frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\Sigma x_i^2}}$

Chương 3. MÔ HÌNH HỒI QUI BỘI

1. Mô hình

Mô hình hồi qui trong đó biến phụ thuộc Y phụ thuộc vào $k - 1$ biến giải thích $X_2, .., X_k$ có dạng

$$E(Y_i) = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \ldots + \beta_k X_{ki} \qquad (1)$$

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \ldots + \beta_k X_{ki} + u_i \qquad (2)$$

Với mẫu $W = \{(X_{2i}, X_{3i}, \ldots, X_{ki}, Y_i);\ i = 1 \div n\}$, SRF có dạng

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \ldots + \hat{\beta}_k X_{ki} \qquad (3)$$

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \ldots + \hat{\beta}_k X_{ki} + e_i \qquad (4)$$

*** Dạng ma trận**

$$\begin{cases} Y_1 = \beta_1 + \beta_2 X_{21} + \ldots + \beta_k X_{k1} + u_1 \\ Y_2 = \beta_1 + \beta_2 X_{22} + \ldots + \beta_k X_{k2} + u_2 \\ \ldots \\ Y_{n-1} = \beta_1 + \beta_2 X_{2n-1} + \ldots + \beta_k X_{kn-1} + u_{n-1} \\ Y_n = \beta_1 + \beta_2 X_{2n} + \ldots + \beta_k X_{kn} + u_n \end{cases} \Leftrightarrow \begin{pmatrix} Y_1 \\ Y_2 \\ \ldots \\ Y_{n-1} \\ Y_n \end{pmatrix} = \begin{pmatrix} 1 & X_{21} & \ldots & X_{k1} \\ 1 & X_{22} & \ldots & X_{k2} \\ \ldots & \ldots & \ldots & \ldots \\ 1 & X_{2n-1} & \ldots & X_{kn-1} \\ 1 & X_{2n} & \ldots & X_{kn} \end{pmatrix} \times \begin{pmatrix} \beta_1 \\ \beta_2 \\ \ldots \\ \beta_k \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \\ \ldots \\ u_{n-1} \\ u_n \end{pmatrix}$$

$$\rightarrow \mathbf{Y}_{(n\times1)} = \mathbf{X}_{(n\times k)} \times \boldsymbol{\beta}_{(k\times1)} + \mathbf{U}_{(n\times1)}$$

$$\mathbf{Y} = \mathbf{X}\times\boldsymbol{\beta} + \mathbf{U} \rightarrow E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$$

Tương tự, đặt $\hat{\mathbf{Y}} = \begin{pmatrix} \hat{Y}_1 \\ \hat{Y}_2 \\ \ldots \\ \hat{Y}_{n-1} \\ \hat{Y}_n \end{pmatrix}$; $\hat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \ldots \\ \hat{\beta}_k \end{pmatrix}$; $\mathbf{e} = \begin{pmatrix} e_1 \\ e_2 \\ \ldots \\ e_{n-1} \\ e_n \end{pmatrix}$, thì $\hat{\mathbf{Y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$

$\mathbf{Y} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e}$

2. Phương pháp bình phương nhỏ nhất

2.1. Phương pháp

Tìm $\hat{\boldsymbol{\beta}}$ sao cho $\sum_{i=1}^{n} e_i^2 = \mathbf{e'e} \rightarrow \min \Leftrightarrow (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \rightarrow \min \Leftrightarrow \mathbf{X'X}\hat{\boldsymbol{\beta}} = \mathbf{X'Y}$

Nếu tồn tại $(\mathbf{X'X})^{-1}$ thì $\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\mathbf{X'Y}$

2.2. Các giả thiết

Gt1 : **X** là phi ngẫu nhiên

Gt2 : $E(\mathbf{U}) = \mathbf{0}$

Gt3 : $Var(u_i) = \sigma^2 \quad \forall i$

Gt4 : $Cov(u_i, u_j) = 0 \quad \forall i \neq j \rightarrow Cov(\mathbf{U}) = \sigma^2\mathbf{I}$ (**I**: ma trận đơn vị)

Gt5 : $Cov(u_i, X_i) = 0 \quad \forall i$

Gt6 : Các biến giải thích không có quan hệ cộng tuyến : $r(\mathbf{X}) = k$

Khi đó $\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\mathbf{X'Y}$ là ước lượng tuyến tính không chệch tốt nhất của $\boldsymbol{\beta}$

2.3. Các tham số của ước lượng

Kì vọng : $E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta}$

Phương sai – hiệp phương sai

$$Cov(\hat{\boldsymbol{\beta}}) = \begin{pmatrix} Var(\hat{\beta}_1) & Cov(\hat{\beta}_1, \hat{\beta}_2) & \dots & Cov(\hat{\beta}_1, \hat{\beta}_k) \\ Cov(\hat{\beta}_2, \hat{\beta}_1) & Var(\hat{\beta}_2) & \dots & Cov(\hat{\beta}_2, \hat{\beta}_k) \\ \dots & \dots & \dots & \dots \\ Cov(\hat{\beta}_k, \hat{\beta}_1) & Cov(\hat{\beta}_k, \hat{\beta}_2) & \dots & Var(\hat{\beta}_k) \end{pmatrix} = \sigma^2(\mathbf{X'X})^{-1}$$

Với σ^2 được ước lượng bởi $\hat{\sigma}^2 = \dfrac{\mathbf{e'e}}{n-k}$

3. Phân tích các hệ số

3.1. Ước lượng khoảng

i.Khoảng tin cậy cho từng hệ số hồi quy

KTC đối xứng : $\hat{\beta}_j - Se(\hat{\beta}_j)t_{\alpha/2}(n-k) < \beta_j < \hat{\beta}_j + Se(\hat{\beta}_j)t_{\alpha/2}(n-k)$

KTC tối đa: $\beta_j < \hat{\beta}_j + Se(\hat{\beta}_j)t_{\alpha}(n-k)$

KTC tối thiểu: $\hat{\beta}_j - Se(\hat{\beta}_j)t_{\alpha}(n-k) < \beta_j$

ii.Khoảng tin cậy cho hai hệ số hồi quy

$$(\hat{\beta}_i \pm \hat{\beta}_j) - Se(\hat{\beta}_i \pm \hat{\beta}_j)t_{\alpha/2}(n-k) < \beta_i \pm \beta_j < (\hat{\beta}_i \pm \hat{\beta}_j) + Se(\hat{\beta}_i \pm \hat{\beta}_j)t_{\alpha/2}(n-k)$$

Với $Se(\hat{\beta}_i \pm \hat{\beta}_j) = \sqrt{Var(\hat{\beta}_i \pm \hat{\beta}_j)} = \sqrt{Var(\hat{\beta}_i) \pm 2Cov(\hat{\beta}_i, \hat{\beta}_j) + Var(\hat{\beta}_j)}$

iii.Khoảng tin cậy cho phương sai yếu tố ngẫu nhiên

KTC 2 phía: $\dfrac{\hat{\sigma}^2(n-k)}{\chi^2_{\alpha/2}(n-k)} < \sigma^2 < \dfrac{\hat{\sigma}^2(n-k)}{\chi^2_{1-\alpha/2}(n-k)}$

KTC tối đa $\sigma^2 < \dfrac{\hat{\sigma}^2(n-k)}{\chi^2_{1-\alpha}(n-k)}$

KTC tối thiểu $\dfrac{\hat{\sigma}^2(n-k)}{\chi^2_{\alpha}(n-k)} < \sigma^2$

3.2. Kiểm định giả thiết

i.Kiểm định giả thiết cho các hệ số hồi quy

Cặp giả thiết	Tiêu chuẩn kiểm định	Miền bác bỏ H_0
$\begin{cases} H_0 : \beta_j = \beta_j^* \\ H_1 : \beta_j \neq \beta_j^* \end{cases}$	$T_{qs} = \dfrac{\hat{\beta}_j - \beta_j^*}{Se(\hat{\beta}_j)}$	$\lvert T_{qs} \rvert > t_{\alpha/2}(n-k)$
$\begin{cases} H_0 : \beta_j = \beta_j^* \\ H_1 : \beta_j > \beta_j^* \end{cases}$		$T_{qs} > t_{\alpha}(n-k)$

$\begin{cases} H_0: \beta_j = \beta_j^* \\ H_1: \beta_j < \beta_j^* \end{cases}$		$T_{qs} < -\ t_\alpha(n-k)$
$\begin{cases} H_0: \beta_i \pm \beta_j = a \\ H_1: \beta_i \pm \beta_j \neq a \end{cases}$	$T_{qs} = \dfrac{\hat{\beta}_i \pm \hat{\beta}_j - a}{Se(\hat{\beta}_i \pm \hat{\beta}_j)}$	$\lvert T_{qs} \rvert > t_{\alpha/2}(n-k)$

ii.Kiểm định giả thiết cho phương sai yếu tố ngẫu nhiên

Cặp giả thiết	Tiêu chuẩn kiểm định	Miền bác bỏ H_0
$\begin{cases} H_0: \sigma^2 = \sigma_0^2 \\ H_1: \sigma^2 \neq \sigma_0^2 \end{cases}$		$\left[\begin{array}{l} \chi^2_{qs} < \chi^2_{1-\alpha/2}(n-k) \\ \chi^2_{qs} > \chi^2_{\alpha/2}(n-k) \end{array}\right.$
$\begin{cases} H_0: \sigma^2 = \sigma_0^2 \\ H_1: \sigma^2 > \sigma_0^2 \end{cases}$	$\chi^2_{qs} = \dfrac{(n-k)\hat{\sigma}^2}{\sigma_0^2}$	$\chi^2_{qs} > \chi^2_\alpha(n-k)$
$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 < \sigma_0^2$		$\chi^2_{qs} < \chi^2_{1-\alpha}(n-k)$

4. Sự phù hợp của hàm hồi qui

4.1. Hệ số xác định

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

Cho biết tỉ lệ sự biến động của biến phụ thuộc được giải thích bởi **tất cả** các biến giải thích có trong mô hình.

Hệ số xác định bội điều chỉnh

$$\overline{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-k} \qquad\qquad \overline{R}^2 < R^2$$

4.2. Kiểm định sự phù hợp của hàm hồi qui

$$\begin{cases} H_0: R^2 = 0 \\ H_1: R^2 \neq 0 \end{cases} \Leftrightarrow \begin{cases} H_0: \beta_2 = \ldots = \beta_k = 0 \\ H_1: \exists \beta_j \neq 0 : (j \neq 1) \end{cases}$$

$$F_{qs} = \frac{ESS/(k-1)}{RSS/(n-k)} = \frac{R^2}{1-R^2} \times \frac{n-k}{k-1}$$

$F_{qs} > F_\alpha(k - 1; n - k)$ thì bác bỏ H_0 : hàm hồi qui là phù hợp

4.3. Kiểm định thu hẹp hồi qui

Nghi ngờ m biến giải thích $X_{k-m+1},\ldots, X_k$ không giải thích cho Y

$$\begin{cases} H_0: \beta_{k-m+1} = \beta_{k-m+2} \ldots = \beta_k = 0 \\ H_1: \exists \beta_j \neq 0 : (j = k-m+1 \div k) \end{cases}$$

$$E(Y/X_2,..,X_{k-m},..,X_k) = \beta_1 + \beta_2 X_2 + \ldots + \beta_k X_k \qquad \text{(L)}$$

$E(Y/X_2,\ldots, X_{k-m}) = \beta_1 + \beta_2 X_2 + \ldots + \beta_k X_{k-m}$ (N)

$$\boxed{F_{qs} = \frac{RSS_N - RSS_L}{RSS_L} \times \frac{n-k}{m} = \frac{R_L^2 - R_N^2}{1 - R_L^2} \times \frac{n-k}{m}}$$

$F_{qs} > F_\alpha(m, n-k)$ bác bỏ H_0

- Trường hợp $m = 1$: $F_{qs} = (T_{qs})^2$ với T_{qs} ứng với hệ số duy nhất cần kiểm định.
- Trường hợp $m = k - 1$: F_{qs} trong kiểm định thu hẹp chính là F_{qs} trong kiểm định sự phù hợp.

5. Dự báo

i. Dự báo giá trị trung bình

$$\hat{Y}_0 - Se(\hat{Y}_0)t_{\alpha/2}(n-k) < E(Y/\mathbf{X}^0) < \hat{Y}_0 + Se(\hat{Y}_0)t_{\alpha/2}(n-k)$$

Với $\hat{Y}_0 = \mathbf{X}^{0\prime}\hat{\boldsymbol{\beta}}$ và $Se(\hat{Y}_0) = \hat{\sigma}\sqrt{\mathbf{X^0}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X^0}}$

ii. Dự báo giá trị cá biệt

$$\hat{Y}_0 - Se(Y_0)t_{\alpha/2}(n-k) < Y_0 < \hat{Y}_0 + Se(Y_0)\, t_{\alpha/2}(n-k)$$

Với $Se(Y_0) = \hat{\sigma}\sqrt{\mathbf{1} + \mathbf{X^0}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X^0}}$

6. Một số mô hình Kinh tế

6.1. Hàm thu nhập – chi tiêu

6.2. Hàm cầu

6.3. Hàm chi phí – sản lượng

6.4. Hàm mũ – Hàm Loga tuyến tính

Mô hình kinh tế có dạng $Y = \beta_0 X_2^{\beta 2} X_3^{\beta 3}$

$$\Leftrightarrow lnY = ln\beta_0 + \beta_2 lnX_2 + \beta_3 lnX_3$$

Xét mô hình $LY = \beta_1 + \beta_2 LX_2 + \beta_3 LX_3 + v$

$$\Leftrightarrow E(Y / X_2 , X_3) = e^{\beta 1} X_2^{\beta 2} X_3^{\beta 3}$$

$\beta_1 : E(Y/X_2 = X_3 = 1) = e^{\beta 1}$

$\beta_2 = \varepsilon_{E(Y)/X2}$: Khi X_2 thay đổi 1%, yếu tố khác không đổi, thì $E(Y)$ thay đổi β_2 %

Ví dụ mô hình : $E(Q) = e^{\beta 1} K^{\beta 2} L^{\beta 3}$

6.5. Hàm chi phí – lợi ích

6.6. Hàm phân tích xu thế

Chương 4. MÔ HÌNH VỚI BIẾN GIẢ

1. Biến định tính – biến giả

1.1. Biến định tính

- Có những yếu tố mang tính **định tính** (*qualitative*) tác động đến biến phụ thuộc
    + Chỉ có một số trạng thái xác định
    + Một cá thể chỉ ở trong một trạng thái, rất khó chuyển sang trạng thái khác
    + Không có đơn vị
- Miêu tả biến định tính bằng biến giả

1.2. Biến giả

VD: Thu nhập có phụ thuộc giới tính ?

Y : thu nhập

$$D = \begin{cases} 1 & \text{Nếu quan sát là Nam} \\ 0 & \text{Nếu quan sát là Nữ} \end{cases}$$

Mô hình : $E(Y/D) = \beta_1 + \beta_2 D$

Thu nhập trung bình của nam $E(Y/D = 1) = \beta_1 + \beta_2$

Thu nhập trung bình của nữ $E(Y/D = 0) = \beta_1$

Nếu $\beta_2 \neq 0$ thì TN trung bình có phụ thuộc giới tính

Biến D đặt như trên là **biến giả** (*dummy variable*).

1.3. Qui tắc đặt biến giả

- Biến giả chỉ nhận giá trị 0 và 1
- Cá thể nào cũng phải có giá trị của biến giả
- Biến giả phân chia tổng thể thành những phần riêng biệt
    → Khi biến định tính có *m* trạng thái

2. Mô hình có biến giải thích chỉ là biến định tính

2.1. Một biến định tính

2.2. Hai biến định tính

VD : Thu nhập trung bình có khác nhau giữa lao động thành thị và nông thôn, nam và nữ?

3. Mô hình có biến giải thích là định tính và định lượng

Xét mô hình tuyến tính Y phụ thuộc vào X có hệ số chặn có dạng:

$$E(Y) = hsc + hsg.X$$

Biến định tính có hai trạng thái A_1 và A_2.

$$D = \begin{cases} 1 & \text{quan sát } \in A_1 \\ 0 & \text{quan sát } \notin A_1 \end{cases}$$

3.1. Biến định tính tác động đến hệ số chặn

$$E(Y/X, D) = \beta_1 + \beta_2 X + \beta_3 D$$

3.2. Biến định tính tác động đến hệ số góc

$$E(Y/X, D) = \beta_1 + \beta_2 X + \beta_3 DX$$

3.3. Tác động đến cả hai hệ số

$$E(Y/X, D) = \beta_1 + \beta_2 X + \beta_3 D + \beta_4 DX$$

$\begin{cases} H_0 : \beta_3 = \beta_4 = 0 \\ H_1 : \beta_3^2 + \beta_4^2 \neq 0 \end{cases}$ Hàm hồi qui đồng nhất trong hai trạng thái

Hàm hồi qui không đồng nhất

3.4. Kiểm định Chow

Kiểm định về sự đồng nhất của hàm hồi qui.

Toàn bộ tổng thể $E(Y) = \beta_1 + \beta_2 X$

Trong A_1 : $E(Y) = \beta_1' + \beta_2' X$

Trong A_2 : $E(Y) = \beta_1'' + \beta_2'' X$

$\begin{cases} H_0 : & [\beta_1' = \beta_1'' = \beta_1] \text{ và } [\beta_2' = \beta_2'' = \beta_2] \\ H_1 : & [\beta_1' \neq \beta_1''] \quad \text{hoặc } [\beta_2' \neq \beta_2''] \end{cases}$

Hàm hồi qui đồng nhất trong hai trạng thái

Hàm hồi qui không đồng nhất

Lấy mẫu W_1 kích thước n_1 trong A_1, hồi qui MH thu được RSS_1

Lấy mẫu W_2 kích thước n_2 trong A_2, hồi qui MH thu được RSS_2

Với mẫu $W = W_1 \cup W_2$ kích thước $n_1 + n_2$, hồi qui thu được RSS

Đặt $\overline{RSS} = RSS_1 + RSS_2$.

$$F_{qs} = \frac{RSS - \overline{RSS}}{\overline{RSS}} \times \frac{n_1 + n_2 - 2k}{k}$$

Nếu $F_{qs} > F_\alpha(k\,;\, n_1 + n_2 - 2k)$: bác bỏ H_0

F_{qs} này và F_{qs} trong kiểm định biến giả sẽ bằng nhau.

4. Hồi qui tuyến tính từng khúc

Hàm hồi qui tuyến tính gấp khúc tại điểm $X = X^*$

$$D = \begin{cases} 1 & : X \geq X^* \\ 0 & : X < X^* \end{cases}$$

$$E(Y/X, D) = \beta_1 + \beta_2 X + \beta_3 (X - X^*) D$$

Chương 5. ĐA CỘNG TUYẾN

1. Hiện tượng đa cộng tuyến

Xét mô hình: $E(Y_i) = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + ... + \beta_k X_{ki}$

GT6: Các biến giải thích không có quan hệ cộng tuyến (mô hình có *từ 2 biến độc lập trở lên*).

Nếu giả thiết bị vi phạm → hiện tượng **đa cộng tuyến** (*Multicollinerity*).

a. Đa cộng tuyến **hoàn hảo** : $\exists\ \lambda_j \neq 0$ ($j \neq 1$) sao cho:

$$\lambda_1 + \lambda_2 X_{2i} + ... + \lambda_k X_{ki} = 0 \quad \forall\ i$$

→ Ma trận X là suy biến, không có lời giải duy nhất.

b. Đa cộng tuyến **không hoàn hảo** : $\exists \lambda_j \neq 0$ ($j \neq 1$) sao cho:

$$\lambda_1 + \lambda_2 X_{2i} + ... + \lambda_k X_{ki} + v_i = 0 \ ,$$

với v_i là YTNN có phương sai dương → vẫn có lời giải.

2. Nguyên nhân

Đa cộng tuyến hoàn hảo gần như không bao giờ xảy ra

Đa cộng tuyến không hoàn hảo thường xuyên xảy ra, do các nguyên nhân:

- Bản chất các biến giải thích có quan hệ hồi qui với nhau.
- Do số liệu mẫu không ngẫu nhiên.
- Do kích thước mẫu không đủ.
- Do quá trình làm trơn số liệu.

3. Hậu quả

Đa cộng tuyến hoàn hảo : không giải được

Đa cộng tuyến không hoàn hảo:

- Các ước lượng có phương sai lớn, là ước lượng không hiệu quả.
- Các kiểm định T có thể sai, khoảng tin cậy rộng không còn ý nghĩa.
- Các ước lượng có thể sai về dấu.
- Kiểm định T và F không thống nhất.

4. Phát hiện

4.1. Sự mâu thuẫn giữa kiểm định T và F

+ Kiểm định F không có ý nghĩa, một kiểm định T về các hệ số góc có ý nghĩa.
+ Kiểm định F có ý nghĩa, tất cả các kiểm định T về các hệ số góc không có ý nghĩa.

$\rightarrow$ có Đa cộng tuyến. Điều ngược lại chưa chắc đúng.

4.2. Hồi qui phụ

Nghi ngờ biến giải thích X_j phụ thuộc tuyến tính vào các biến giải thích khác, dùng mô hình **hồi qui phụ** (*auxilliary regression*)

$$X_j = \alpha_1 + \alpha_2 X_2 + \ldots + \alpha_{j-1} X_{j-1} + \alpha_{j+1} X_{j+1} + \ldots + v \quad (*)$$

$\begin{cases} H_0 : R_*^2 = 0 \\ H_1 : R_*^2 \neq 0 \end{cases}$ Mô hình ban đầu không có Đa cộng tuyến / Mô hình ban đầu có Đa cộng tuyến

- $H_0 : R_*^2 = 0$ — Mô hình ban đầu không có Đa cộng tuyến
- $H_1 : R_*^2 \neq 0$ — Mô hình ban đầu có Đa cộng tuyến

$\rightarrow$ $F_{qs} = \dfrac{R_*^2}{1-R_*^2} \times \dfrac{n-k_*}{k_*-1}$; $F_{qs} > F_\alpha(k_* - 1, n - k_*)$ thì bác bỏ H_0.

Có thể dùng kiểm định T có các hệ số tương ứng.

(* *Có nhiều hồi qui phụ để kiểm định cho hiện tượng Đa cộng tuyến*)

4.3. Độ đo Theil

Dùng để so sánh mức độ đa cộng tuyến không hoàn hảo giữa các mô hình

Khi bỏ biến X_j ra khỏi mô hình, hồi qui thu được R^2_{-j}

$$m = R^2 - \sum_{j=2}^{k}(R^2 - R^2_{-j}) \text{ được gọi là độ đo Theil}$$

5. Khắc phục

- Bỏ bớt biến
- Lấy thêm mẫu
- Đổi dạng của mô hình

Chương 6. PHƯƠNG SAI SAI SỐ THAY ĐỔI

1. Hiện tượng phương sai sai số thay đổi

MH ban đầu: $Y_i = \beta_1 + \beta_2 X_i + u_i$

Gt 3: Phương sai các yếu tố ngẫu nhiên là đồng nhất $Var(u_i) \equiv \sigma^2$ không đổi.

Nếu gt được thỏa mãn → **Phương sai sai số đồng đều** (không đổi - *homoscocedasticity*).

Gt không thỏa mãn : $Var(u_i) = \sigma_i^2$ không đồng nhất → PSSS thay đổi (*heterscocedasticity*).

2. Nguyên nhân

- Bản chất hiện tượng Kinh tế xã hội.
- Số liệu không đúng bản chất hiện tượng.
- Quá trình xử lý số liệu.

3. Hậu quả

- Các ước lượng là không chệch, nhưng không hiệu quả → không phải là tốt nhất.
- Các kiểm định T, F có thể sai, khoảng tin cậy rộng.

4. Phát hiện

$Var(u_i) = \sigma_i^2$ chưa biết. Ta dùng ước lượng của nó là e_i^2 để phân tích đánh giá.

4.1. Đồ thị phần dư

Dùng đồ thị của e_i, $|e_i|$ hoặc e_i^2 để đánh giá.

4.2. Kiểm định Glejer

Gt : $\sigma_i^2 = \sigma^2 X_i$, do đó hồi qui mô hình hồi qui phụ

$$e_i^2 = \alpha_1 + \alpha_2 X_i + v_i \quad (*)$$

$$\begin{cases} H_0 : \alpha_2 = 0 : R_*^2 = 0 & \text{Mô hình đầu có PSSS đồng đều} \\ H_1 : \alpha_2 \neq 0 : R_*^2 \neq 0 & \text{Mô hình đầu có PSSS thay đổi} \end{cases}$$

Dùng kiểm định T hoặc F để kiểm định

Tương tự

Gt : $\sigma_i^2 = \sigma^2 X_i^2$ → MH hồi qui phụ $e_i^2 = \alpha_1 + \alpha_2 X_i^2 + v_i$

Gt : $\sigma_i^2 = \sigma^2 \sqrt{X_i}$ → MH hồi qui phụ $e_i^2 = \alpha_1 + \alpha_2 \sqrt{X_i} + v_i$

Gt : $\sigma_i^2 = \sigma^2 \frac{1}{X_i}$ → MH hồi qui phụ $e_i^2 = \alpha_1 + \alpha_2 \frac{1}{X_i} + v_i$

Có thể sử dụng $|e_i|$ để đại diện cho $Se(u_i)$, mô hình hồi qui phụ sẽ có thay đổi tương ứng.

4.3. Kiểm định Park

Giả thiết: $\sigma_i^2 = \sigma^2 X_i^{\alpha 2}$ → MH hồi qui phụ $ln e_i^2 = \alpha_1 + \alpha_2 ln X_i + v_i$

4.4. Kiểm định White

Dùng cho mô hình nhiều biến giải thích. Hồi qui bình phương phần dư theo tổ hợp bậc cao dần của các biến giải thích.

VD : MH ban đầu $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$

→ MH hồi qui phụ : $e^2 = \alpha_1 + \alpha_2 X_2 + \alpha_3 X3 + \alpha_4 X_2^{\ 2} + \alpha_5 X_3^{\ 2} + \alpha_6 X_2 X_3$ $(+\dots+) + v_i$ (*)

$$\begin{cases} H_0 : R_*^2 = 0 \\ H_1 : R_*^2 \neq 0 \end{cases}$$

Kiểm định χ^2 : $\chi^2_{qs} = nR_*^2$, nếu $\chi^2_{qs} > \chi^2_\alpha(k_* - 1)$ thì bác bỏ H_0

4.5. Kiểm định dựa trên biến phụ thuộc

Giả thiết phương sai sai số **thay đổi theo bình phương trung bình biến phụ thuộc** $\sigma_i^2 = \sigma^2 E(Yi)^2$

B1: Hồi qui mô hình gốc thu được phần dư e_i và giá trị ước lượng $\hat{Y}_i$

B2 : Hồi qui mô hình hồi qui phụ $e_i^2 = \alpha_1 + \alpha_2 \hat{Y}_i^2 + v_i$ (*)

$$\begin{cases} H_0 : \alpha_2 = 0 \\ H_1 : \alpha_2 \neq 0 \end{cases} \Leftrightarrow \begin{cases} H_0 : R_*^2 = 0 \\ H_1 : R_*^2 \neq 0 \end{cases}$$ Kiểm định T, F, χ^2

Kiểm định χ^2 : $\chi^2_{qs} = nR_*^2$, nếu $\chi^2_{qs} > \chi^2_\alpha(1)$ thì bác bỏ H_0

5. Khắc phục

Dựa trên giả thiết về sự thay đổi của PSSS thay đổi mà khắc phục

5.1. Nếu biết σ_i^2

Chia hai vế mô hình cho σ_i

$$\frac{Y_i}{\sigma_i} = \beta_1 \frac{1}{\sigma_i} + \beta_2 \frac{X_i}{\sigma_i} + \frac{u_i}{\sigma_i} \quad \Leftrightarrow \quad Y_i' = \beta_1 X_{0i} + \beta_2 X_i' + u_i'$$

$Var(u_i') = 1$ không đổi

5.2. Nếu chưa biết σ_i^2

Gt : $\sigma_i^2 = \sigma^2 X_i$: chia hai vế cho $\sqrt{X_i}$

$$\frac{Y_i}{\sqrt{X_i}} = \beta_1 \frac{1}{\sqrt{X_i}} + \beta_2 \sqrt{X_i} + \frac{u_i}{\sqrt{X_i}}$$ PSSS sẽ bằng σ^2

Gt : $\sigma_i^2 = \sigma^2 X_i^2$: chia hai vế cho X_i

Gt : $\sigma_i^2 = \sigma^2 E(Yi)^2$: chia hai vế cho $\hat{Y}_i$

Chương 7. TỰ TƯƠNG QUAN

1.Hiện tượng tự tương quan

Mô hình ban đầu: $Y_t = \beta_1 + \beta_2 X_t + u_t$

Gt 4: Các yếu tố ngẫu nhiên không tương quan

$Cov(u_i, u_j) = 0$ $(i \neq j)$ hoặc $Cov(u_t, u_{t-p}) = 0$ $(p \neq 0)$

Nếu gt bị vi phạm : hiện tượng tự tương quan bậc p (*Autocorrelation order p*)

Xét trường hợp $p = 1$

u_t và u_{t-1} có cùng trung bình và phương sai

$\rightarrow$ $u_t = \rho u_{t-1} + \varepsilon_t$ ($-1 \leq \rho \leq 1$, ε_t thỏa mãn các giả thiết của OLS)

- $\rho = -1$ tự tương quan âm hoàn toàn
- $-1 < \rho < 0$ tự tương quan âm
- $\rho = 0$ không có tự tương quan
- $0 < \rho < 1$ tự tương quan dương
- $\rho = 1$ tự tương quan dương hoàn toàn

Tự tương quan bậc p : $u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \ldots + \rho_p u_{t-p} + \varepsilon_t$ với $\rho_p \neq 0$

2. Nguyên nhân

- Bản chất, tính quán tính trong hiện tượng kinh tế xã hội
- Hiện tượng mạng nhện trong kinh tế
- Quá trình xử lý, nội suy số liệu
- Mô hình thiếu biến hoặc dạng hàm sai

3. Hậu quả

Các ước lượng là không chệch nhưng không còn là ước lượng tốt nhất.

4. Phát hiện

4.1. Kiểm định Durbin – Watson

Dùng để kiểm định tự tương quan bậc 1 : $u_t = \rho u_{t-1} + \varepsilon_t$

$$d = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2} \cong 2(1 - \hat{\rho}) \text{ với } \hat{\rho} = \frac{\sum_{i=1}^{n} e_t e_{t-1}}{\sum_{i=1}^{n} e_t^2}$$ là ước lượng cho ρ

$-1 \leq \hat{\rho} \leq 1 \rightarrow 0 \leq d \leq 4$

Với n, $k' = k - 1$ cho trước, tra bảng $\rightarrow d_L$ và d_U

Tự tương quan dương $\rho > 0$	Không có kết luận	Không có tự tương quan $\rho = 0$	Không có kết luận	Tự tương quan âm $\rho < 0$

0 d_L d_U 2 $4 - d_U$ $4 - d_L$

Chú ý. Kiểm định DW sẽ không dùng được khi mô hình không có hệ số chặn, hoặc có trễ bậc một của biến phụ thuộc làm biến giải thích.
Trường hợp mô hình có trễ bậc 1 của biến phụ thuộc làm biến giải thích thì dùng Durbin-Watson h:

Ví dụ : $Y_t = \beta_1 + \beta_2 X_t + \alpha_1 Y_{t-1} + u_t$ hay $Y = \beta_1 + \beta_2 X + \alpha_1 Y(-1) + u$

$$h = \hat{\rho}\sqrt{\frac{n}{1 - nVar(\hat{\alpha}_1)}}\ ;\ \hat{\rho} = 1 - \frac{d}{2}$$

Tự tương quan âm $\rho < 0$	Không có tự tương quan $\rho = 0$	Tự tương quan dương $\rho > 0$

- 1.96 1.96

4.2. Hồi qui phụ

Kiểm định

$$u_t = \rho u_{t-1} + \rho u_{t-2} + \ldots + \rho u_{t-p} + \varepsilon_t$$

Mô hình hồi qui phụ : $e_t = (\alpha_0) + \alpha_1 e_{t-1} + \ldots + \alpha_p e_{t-p} + v_t$ (*)

$\begin{cases} H_0 : R_*^2 = 0 \\ H_1 : R_*^2 \neq 0 \end{cases} \Leftrightarrow \begin{cases} H_0 : \alpha_1 = \ldots = \alpha_p = 0 \\ H_1 : \exists \alpha_j \neq 0 : (j \neq 0) \end{cases}$ Không có tự tương quan đến bậc p; Có tự tương quan ở bậc tương ứng

Kiểm định T hoặc F

4.3. Kiểm định BG

Mô hình hồi qui phụ

$$e_t = [\beta_1 + \beta_2 X_t] + \alpha_1 e_{t-1} + \ldots + \alpha_p e_{t-p} + v_t \ (*)$$

$H_0 : \alpha_1 = \ldots = \alpha_p = 0$

$H_1 : \exists \alpha_j \neq 0 : (j \neq 0)$

Kiểm định χ^2 : $\chi^2_{qs} = n_* R_*^2 = (n-p)R_*^2$, nếu $\chi^2_{qs} > \chi^2_\alpha(p)$ thì bác bỏ H_0

Kiểm định F:

Hồi qui

$e_t = [\beta_1 + \beta_2 X_t] + v_t$ (**)

$F_{qs} = \dfrac{R_*^2 - R_{**}^2}{1 - R_*^2} \times \dfrac{n_* - k_*}{k_* - 1}$ Nếu $F_{qs} > F_\alpha(k_* - 1; n_* - k_*)$ thì bác bỏ H_0

5. Khắc phục

Mục đích là chuyển mô hình ban đầu có khuyết tật tự tương quan thành mô hình mới không có tự tương quan.

Mô hình ban đầu: $Y_t = \beta_1 + \beta_2 X_t + u_t$

Có tự tương quan : $u_t = \rho u_{t-1} + \varepsilon_t$ với $\rho \neq 0$, ε_t thỏa mãn các giả thiết OLS.

5.1. Khi ρ đã biết

$Y_t = \beta_1 + \beta_2 X_t + u_t$

$\leftrightarrow Y_{t-1} = \beta_1 + \beta_2 X_{t-1} + u_{t-1} \leftrightarrow \rho Y_{t-1} = \rho\beta_1 + \beta_2 \rho X_{t-1} + \rho u_{t-1}$

$\rightarrow Y_t - \rho Y_{t-1} = \beta_1(1-\rho) + \beta_2(X_t - \rho X_{t-1}) + u_t - \rho u_{t-1}$ (phương trình sai tổng quát)

$\rightarrow \quad Y_t^* = \beta_1^* + \beta_2 X_t^* + \varepsilon_t$

Ước lượng bằng OLS $\rightarrow \hat{\beta}_1^* \rightarrow \hat{\beta}_1 = \dfrac{\hat{\beta}_1^*}{1-\rho}$ và $\hat{\beta}_2$

a. Trường hợp tự tương quan dương hoàn toàn $\rho = 1$

PT sai phân tổng quát $\Leftrightarrow \Delta Y_t = \beta_2 \Delta X_t + \varepsilon_t$ (phương trình sai phân cấp 1)

b. Trường hợp tự tương quan âm hoàn toàn $\rho = -1$

PT sai phân tổng quát $\Leftrightarrow \dfrac{Y_t + Y_{t-1}}{2} = \beta_1 + \beta_2 \dfrac{X_t + X_{t-1}}{2} + \varepsilon_t$ (mô hình trung bình trượt)

5.2. Khi ρ chưa biết

Ước lượng ρ bằng các phương pháp khác nhau để thay vào phương trình sai phân tổng quát

Từ thống kê Durbin-Watson

$d \cong 2(1 - \hat{\rho}) \rightarrow \hat{\rho} = 1 - \dfrac{d}{2}$

Từ hồi qui phụ

$e_t = (\alpha_0) + \alpha_1 e_{t-1} + v_t \rightarrow$ lấy $\hat{\rho} = \hat{\alpha}_1$

Phương pháp Cochran-Orcutt

Hồi qui mô hình ban đầu: $Y_t = \beta_1 + \beta_2 X_t + u_t \rightarrow \hat{\beta}_1^{(1)}, \hat{\beta}_2^{(1)}, e_t^{(1)}$

Hồi qui mô hình $e_t = \alpha_0 + \alpha_1 e_{t-1} + v_t \rightarrow \hat{\alpha}_1^{(1)}$

Lấy $\hat{\rho} = \hat{\alpha}_1^{(1)}$ thay vào phương trình sai phân tổng quát $\rightarrow \hat{\beta}_1^{(2)}, \hat{\beta}_2^{(2)}, e_t^{(2)}$

Hồi qui mô hình $e_t = \alpha_0 + \alpha_1 e_{t-1} + v_t \rightarrow \hat{\alpha}_1^{(2)}$

Lấy $\hat{\rho} = \hat{\alpha}_1^{(2)}$

…

Quá trình lặp cho đến khi $\hat{\rho}$ ở hai bước kế tiếp chênh lệch nhau không đáng kể, $\hat{\beta}_1$ và $\hat{\beta}_2$ ở bước cuối cùng là ước lượng cho β_1 và β_2.

Chương 8. ĐỊNH DẠNG MÔ HÌNH

1. Thuộc tính của mô hình tốt

- Đầy đủ
- Phù hợp
- Khả năng phân tích và dự báo

2. Mô hình thừa biến giải thích

Nếu mô hình thừa biến giải thích thì các ước lượng vẫn là không chệch và vững, nhưng không hiệu quả, khoảng tin cậy rộng.

Kiểm định bằng cách bỏ bớt biến số nghi là không cần thiết và dùng kiểm định với hệ số tương ứng để kết luận

3. Mô hình thiếu biến – dạng hàm sai

Xét mô hình ban đầu : $Y = \beta_1 + \beta_2 X + u$ (1)

a. Kiểm định Ramsey

B1: Hồi qui mô hình ban đầu thu được các giá trị ước lượng $\hat{Y}$

B2: Hồi qui mô hình hồi qui phụ :

$$Y = [\beta_1 + \beta_2 X] + \alpha_1 \hat{Y}^2 + \ldots + \alpha_m \hat{Y}^{m+1} + u \quad (2)$$

$\mathrm{H}_0 : \alpha_1 = \ldots = \alpha_m = 0$ MH (1) không thiếu biến

$\mathrm{H}_1 : \exists \alpha_j \neq 0, j = \overline{1, m}$ MH (1) thiếu biến

$$F_{qs} = \frac{R^2_{(2)} - R^2_{(1)}}{1 - R^2_{(2)}} \times \frac{n - k_{(2)}}{k_{(2)} - 1}$$

Nếu $F_{qs} > F_\alpha(k_{(2)} - 1; n - k_{(2)})$ bác bỏ H_0

b. Kiểm định nhân tử Lagrange (LM)

B1: Hồi qui mô hình ban đầu thu được các phần dư *e* và giá trị ước lượng $\hat{Y}$

B2: Hồi qui mô hình hồi qui phụ :

$$e = [\beta_1 + \beta_2 X] + \alpha_1 \hat{Y}^2 + \ldots + \alpha_m \hat{Y}^{m+1} + v \quad (*)$$

$\mathrm{H}_0 : \alpha_1 = \ldots = \alpha_m = 0$ MH (1) có dạng hàm đúng

$\mathrm{H}_1 : \exists \alpha_j \neq 0, j = \overline{1, m}$ MH (1) có dạng hàm sai

Kiểm định χ^2 : $\chi^2_{qs} = nR^2_*$, nếu $\chi^2_{qs} > \chi^2_\alpha(m)$ thì bác bỏ H_0.

4. Phân phối xác suất yếu tố ngẫu nhiên

Các phân tích về hệ số (khoảng tin cậy, kiểm định giả thiết) phụ thuộc giả thiết YTNN phân phối chuẩn. Nếu YTNN không phân phối chuẩn thì các ước lượng vẫn là ước lượng tốt nhất, nhưng các phân tích không dùng được.

H_0 : YTNN phân phối chuẩn (*Normality distribution*)

H_1 : YTNN không phân phối chuẩn

Kiểm định χ^2 : JB = $\chi^2_{qs} = n\left[\dfrac{S^2}{6} + \dfrac{(K-3)^2}{24}\right]$

Với *S* là hệ số bất đối xứng (*skewness*), *K* là hệ số nhọn (*kutosis*)

Nếu $\chi_{qs}^2 > \chi_{\alpha}^2(2)$ thì bác bỏ H_0