



CHƯƠNG 4 – HỒI QUY THEO BIẾN GIÁ

Nội dung

4.1. Bản chất của biến giả (Dummy variable)

4.2. Hồi quy với biến giả

4.3. Ứng dụng của biến giả

4.1. Bản chất của biến giả (Dummy variable)

- **Biến định tính hay biến chất lượng là biến phân loại các tính chất, phạm trù khác nhau.**
- **Ví dụ 1:**
 - Biến giới tính: có hai phạm trù là nam và nữ
 - Biến vùng miền: có ba phạm trù là miền bắc, miền trung và miền nam
- **Bản chất của biến giả:**
 - Thường được ký hiệu là D (Dummy Variable)
 - Là kỹ thuật được dùng để lượng hóa các biến chất lượng.
 - Thông thường biến giả chỉ có hai giá trị là 0 và 1 nên còn được gọi là biến nhị phân.

4.1. Bản chất của biến giả (Dummy variable)

■ Nguyên tắc lượng hóa biến định tính thông qua kỹ thuật biến giả:

- Để lượng hóa biến định tính có hai phạm trù: dùng một biến giả.
- Để lượng hóa biến định tính có m phạm trù: dùng $(m-1)$ biến giả.

Trong đó:

- Phạm trù ứng với biến giả nhận giá trị bằng 1 được gọi là phạm trù so sánh.
- Phạm trù ứng với biến giả nhận giá trị bằng 0 được gọi là phạm trù cơ sở.

4.2. Hồi quy với biến giả

4.2.1. Mô hình có một biến độc lập là biến định tính (có hai phạm trù)

■ **Ví dụ 2:** Hồi quy thu nhập của công chức (Y) phụ thuộc vào giới tính, giả sử biến giới tính có hai phạm trù là nam và nữ:

- Sử dụng biến giả D để lượng hóa biến giới tính như sau:

$$D_i = \begin{cases} 1 & \text{nếu công chức là nam} \\ 0 & \text{nếu công chức là nữ} \end{cases}$$

- Khi đó, hàm hồi quy tổng thể có dạng:

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

4.2. Hồi quy với biến giả

Hàm hồi quy tổng thể tại các phạm trù có dạng như sau:

- Thu nhập trung bình của công chức nữ:

$$E(Y / D_i = 0) = \beta_1$$

- Thu nhập trung bình của công chức nam:

$$E(Y / D_i = 1) = \beta_1 + \beta_2$$

- Để xem xét có sự phân biệt về giới tính trong thu nhập hay không, thực hiện kiểm định các cặp giả thuyết:

$$(1) \begin{cases} H_0 : \beta_2 = 0 \\ H_1 : \beta_2 \neq 0 \end{cases}$$

$$(2) \begin{cases} H_0 : \beta_2 = 0 \\ H_1 : \beta_2 > 0 \end{cases}$$

4.2. Hồi quy với biến giả

- Ví dụ 3:** Giả sử có số liệu về mức lương khởi điểm của giáo sư đại học Y (triệu đồng) theo giới tính với $D1 = 0$ (nữ), $D1 = 1$ (nam). Hồi quy mô hình thu được:

Bảng 4.1. Kết quả hồi quy mức lương theo giới tính

Y	D1
22	1
19	0
18	0
21.7	1
18.5	0
21	1
20.5	1
17	0
17.5	0
21.2	1

Equation: UNTITLED Workfile: UNTITLED::Untitled\

ViewProcObjectPrintNameFreezeEstimateForecastStatsResids

Dependent Variable: Y
Method: Least Squares
Date: 01/03/19 Time: 15:19
Sample: 1 10
Included observations: 10

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D1	3.280000	0.440908	7.439191	0.0001
C	18.00000	0.311769	57.73503	0.0000
R-squared	0.873701	Mean dependent var	19.64000	
Adjusted R-squared	0.857913	S.D. dependent var	1.849444	
S.E. of regression	0.697137	Akaike info criterion	2.293187	
Sum squared resid	3.888000	Schwarz criterion	2.353704	
Log likelihood	-9.465934	Hannan-Quinn criter.	2.226800	
F-statistic	55.34156	Durbin-Watson stat	0.667284	
Prob(F-statistic)	0.000073			

4.2. Hồi quy với biến giả

- Hàm hồi quy mẫu có dạng:

$$\hat{Y}_i = 18 + 3.28D1_i$$

$$E(Y / D1 = 0) = 18$$

$$E(Y / D1 = 1) = 18 + 3.28 = 21.28$$

- Từ kết quả hồi quy cho ta thấy: Mức lương trung bình của GS nữ là 18 triệu đồng, trong khi đó mức lương trung bình của GS nam là 21.28 triệu đồng.
- Giá trị $\hat{\beta}_2 = 3.28$ cho ta biết chênh lệch mức lương trung bình của GS nam so với GS nữ.

4.2. Hồi quy với biến giả

➤ Chú ý: Nếu sử dụng hai biến giả để lượng hóa cho biến định tính có hai phạm trù sẽ vi phạm giả thiết của OLS.

• Xét mô hình:

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 D_{2i} + \hat{\beta}_3 D_{3i} + U_i$$

Trong đó

$$D_{2i} = \begin{cases} 1 & \text{nếu là nam giáo sư} \\ 0 & \text{nếu không là nam giáo sư} \end{cases}$$
$$D_{3i} = \begin{cases} 1 & \text{nếu là nữ giáo sư} \\ 0 & \text{nếu không là nữ giáo sư} \end{cases}$$

Khi đó $D_2 + D_3 = 1$ (véc tơ có mọi thành phần bằng 1). Do đó ma trận XT X của mô hình hồi quy tuyến tính suy biến, vi phạm giả thiết các biến độc lập.

4.2. Hồi quy với biến giả

4.2.1. Mô hình có một biến độc lập là biến định tính (có nhiều phạm trù)

- **Ví dụ 4:** Hồi quy thu nhập của công chức (Y) phụ thuộc vào khu vực làm việc. Trong đó, biến khu vực làm việc có 3 vùng: nông thôn; thành thị và miền núi. Khi đó, sử dụng 2 biến giả D_2 và D_3 để lượng hóa biến khu vực làm việc như sau:

$$D_{2i} = \begin{cases} 1 & \text{nếu công chức } i \text{ làm việc ở nông thôn} \\ 0 & \text{nếu công chức } i \text{ làm việc ở khu vực khác} \end{cases}$$

$$D_{3i} = \begin{cases} 1 & \text{nếu công chức } i \text{ làm việc ở thành thị} \\ 0 & \text{nếu công chức } i \text{ làm việc ở khu vực khác} \end{cases}$$

- Hàm hồi quy tổng thể: $E(Y / D_2, D_3) = \beta_1 + \beta_2 D_{2i} + \beta_3 D_{3i}$

4.2. Hồi quy với biến giả

- Thu nhập trung bình của công chức làm việc ở miền núi:

$$E(Y / D_{2i} = D_{3i} = 0) = \beta_1$$

- Thu nhập trung bình của công chức làm việc ở nông thôn:

$$E(Y / D_{2i} = 1, D_{3i} = 0) = \beta_1 + \beta_2$$

- Thu nhập trung bình của công chức làm việc ở thành thị:

$$E(Y / D_{2i} = 0, D_{3i} = 1) = \beta_1 + \beta_3$$

- Để xem có sự khác biệt về thu nhập giữa công chức ở các khu vực hay không ta kiểm định các cặp giả thuyết:

$$(1) \begin{cases} H_0 : \beta_j = 0 \\ H_1 : \beta_j \neq 0 \end{cases} (j = 2, 3)$$

$$(2) \begin{cases} H_0 : \beta_2 = \beta_3 \\ H_1 : \beta_2 \neq \beta_3 \end{cases}$$

4.2. Hồi quy với biến giả

4.2.2. Mô hình có hai biến độc lập là biến định tính

Tùy thuộc vào số phạm trù của mỗi biến định tính để sử dụng số biến giả thích hợp trong mô hình.

- **Ví dụ 5:** Hồi quy thu nhập của công chức (Y) phụ thuộc vào giới tính và khu vực làm việc. Trong đó, biến giới tính có hai phạm trù là nam và nữ; biến khu vực làm việc có ba phạm trù là nông thôn, thành thị và miền núi. Khi đó, sử dụng 3 biến giả D_2, D_3, D_4 .

- Mô hình hồi quy:

$$Y_i = \beta_1 + \beta_2 D_{2i} + \beta_3 D_{3i} + \beta_4 D_{4i} + U_i$$

4.2. Hồi quy với biến giả

Trong đó:

$$D_{2i} = \begin{cases} 1 & \text{nếu công chức là nam} \\ 0 & \text{nếu công chức là nữ} \end{cases}$$

$$D_{3i} = \begin{cases} 1 & \text{nếu công chức i làm việc ở nông thôn} \\ 0 & \text{nếu công chức i làm việc ở khu vực khác} \end{cases}$$

$$D_{4i} = \begin{cases} 1 & \text{nếu công chức i làm việc ở thành thị} \\ 0 & \text{nếu công chức i làm việc ở khu vực khác} \end{cases}$$

4.2. Hồi quy với biến giả

4.2.3. Mô hình có một biến định lượng và một biến định tính (có hai phạm trù)

- **Ví dụ 6:** Hồi quy lương trung bình của một giáo sư đại học (Y) phụ thuộc vào số năm kinh nghiệm giảng dạy (X) và giới tính với hai phạm trù nam và nữ. Sử dụng biến giả D lượng hóa biến giới tính như sau:

$$D_i = \begin{cases} 1 & \text{nếu là giáo sư nam} \\ 0 & \text{nếu là giáo sư nữ} \end{cases}$$

Hàm hồi quy tổng thể: $E(Y / X_i, D_i) = \beta_1 + \beta_2 X_i + \beta_3 D_i$

- Mức lương trung bình của giáo sư nữ: $E(Y_i / X_i, D_i = 0) = (\beta_1) + \beta_2 X_i$
- Mức lương trung bình của giáo sư nam: $E(Y_i / X_i, D_i = 1) = (\beta_1 + \beta_3) + \beta_2 X_i$

4.2. Hồi quy với biến giả

4.2.3. Mô hình có một biến định lượng và một biến định tính (có nhiều phạm trù)

- **Ví dụ 7:** Nghiên cứu chi tiêu cá nhân cho dịch vụ y tế hàng năm (Y), phụ thuộc vào thu nhập (X) và trình độ học vấn với ba cấp: dưới trung học, trung học và đại học. Sử dụng hai biến giả D_1 và D_2 để lượng hóa biến trình độ học vấn như sau:

$$D_{1i} = \begin{cases} 1 & \text{Có trình độ đại học} \\ 0 & \text{Trình độ khác} \end{cases}$$

$$D_{2i} = \begin{cases} 1 & \text{Có trình độ trung học} \\ 0 & \text{Trình độ khác} \end{cases}$$

4.2. Hồi quy với biến giả

Hàm hồi quy tổng thể:

$$E(Y / X_i, D_{1i}, D_{2i}) = \beta_1 + \beta_2 X_i + \beta_3 D_{1i} + \beta_4 D_{2i}$$

- Chi tiêu cá nhân của trình độ dưới trung học:

$$E(Y_i / X_i, D_{1i} = 0, D_{2i} = 0) = (\beta_1) + \beta_2 X_i$$

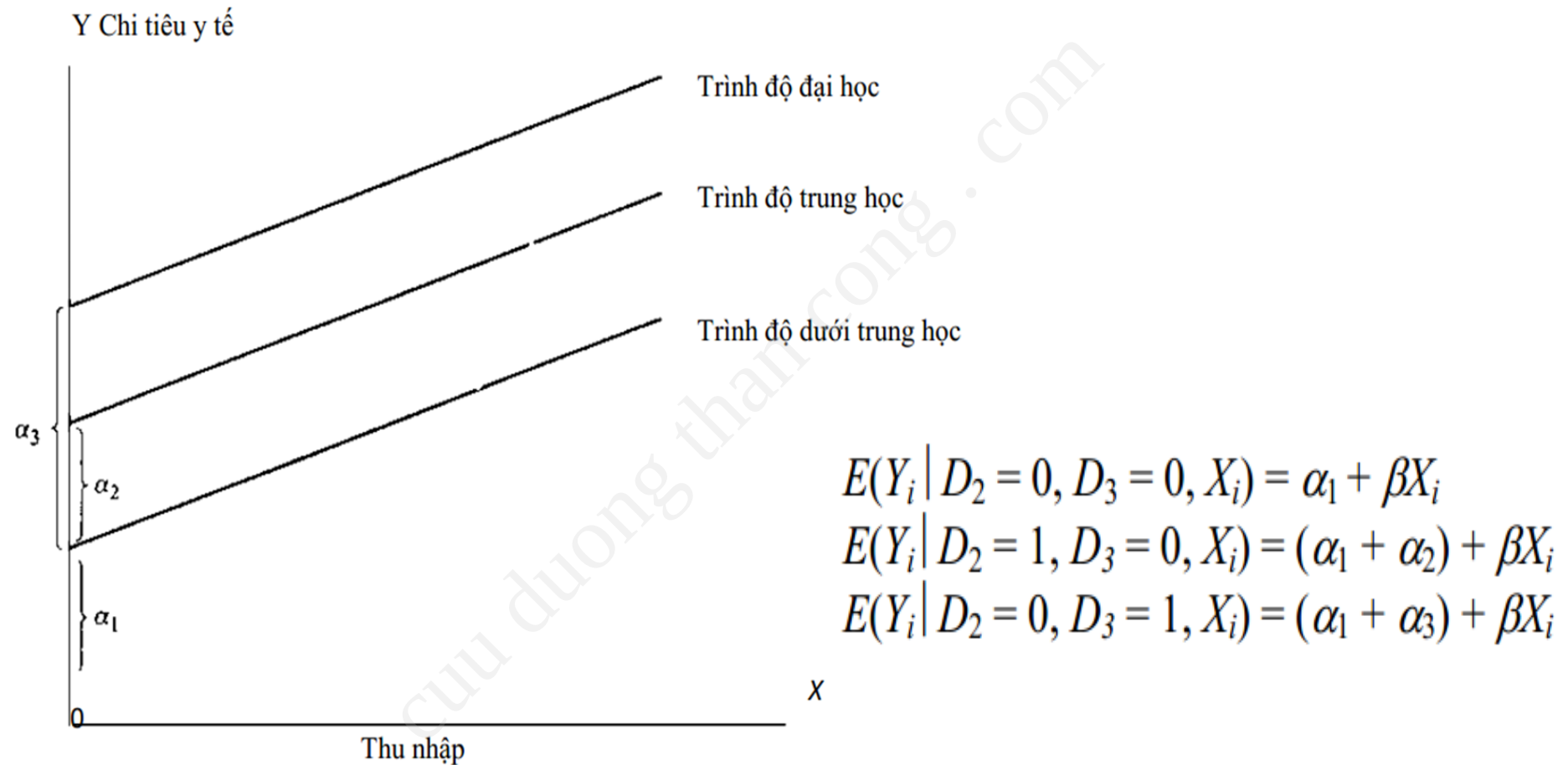
- Chi tiêu cá nhân của trình độ trung học:

$$E(Y_i / X_i, D_{1i} = 1, D_{2i} = 0) = (\beta_1 + \beta_3) + \beta_2 X_i$$

- Chi tiêu cá nhân của trình độ đại học:

$$E(Y_i / X_i, D_{2i} = 1, D_{1i} = 0) = (\beta_1 + \beta_4) + \beta_2 X_i$$

4.2. Hồi quy với biến giả



Hình 4.1. Đường hồi quy tổng thể

4.2. Hồi quy với biến giả

4.2.4. Mô hình có nhiều biến định lượng và nhiều biến định tính

- **Ví dụ 8:** Hồi quy lương trung bình của một công nhân (Y) phụ thuộc vào số năm kinh nghiệm (X_1), trình độ học vấn (X_2), giới tính (với hai phạm trù nam và nữ) và vùng miền (với ba phạm trù miền Bắc, miền Nam và miền Trung). Sử dụng ba biến giả D_1 , D_2 , D_3 như sau:

$$D_{1i} = \begin{cases} 1 & \text{Nếu là công nhân nam} \\ 0 & \text{Nếu là công nhân nữ} \end{cases}$$

$$D_{2i} = \begin{cases} 1 & \text{nếu là người miền Bắc} \\ 0 & \text{nếu không là người miền Bắc} \end{cases} \quad D_{3i} = \begin{cases} 1 & \text{nếu là người miền Nam} \\ 0 & \text{nếu không là người miền Nam} \end{cases}$$

4.2. Hồi quy với biến giả

- Hàm hồi quy tổng thể:

$$E(Y / X_{2i}, X_{3i}, \dots) = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 D_{1i} + \beta_5 D_{2i} + \beta_6 D_{3i}$$

- Tiền lương trung bình của công nhân nữ ở miền Trung:

$$E(Y_i / X_i, D_{1i} = 0, D_{2i} = 0, D_{3i} = 0) = (\beta_1) + \beta_2 X_{2i} + \beta_3 X_{3i}$$

- Tiền lương trung bình của công nhân nữ ở miền Bắc:

$$E(Y_i / X_i, D_{1i} = 0, D_{2i} = 1, D_{3i} = 0) = (\beta_1 + \beta_5) + \beta_2 X_{2i} + \beta_3 X_{3i}$$

- Tiền lương trung bình của công nhân nữ ở miền Nam:

$$E(Y_i / X_i, D_{1i} = 0, D_{2i} = 0, D_{3i} = 1) = (\beta_1 + \beta_6) + \beta_2 X_{2i} + \beta_3 X_{3i}$$

4.2. Hồi quy với biến giả

- Tiền lương trung bình của công nhân nam ở miền Trung:

$$E(Y_i / X_i, D_{1i} = 1, D_{2i} = 0, D_{3i} = 0) = (\beta_1 + \beta_4) + \beta_2 X_{2i} + \beta_3 X_{3i}$$

- Tiền lương trung bình của công nhân nam ở miền Bắc:

$$E(Y_i / X_i, D_{1i} = 1, D_{2i} = 1, D_{3i} = 0) = (\beta_1 + \beta_4 + \beta_5) + \beta_2 X_{2i} + \beta_3 X_{3i}$$

- Tiền lương trung bình của công nhân nam ở miền Nam:

$$E(Y_i / X_i, D_{1i} = 1, D_{2i} = 0, D_{3i} = 1) = (\beta_1 + \beta_4 + \beta_6) + \beta_2 X_{2i} + \beta_3 X_{3i}$$

4.2. Hồi quy với biến giả

4.2.5. Phân tích ảnh hưởng tương tác giữa các biến định tính

Trong các mô hình hồi quy có nhiều biến giả có thể xảy ra sự tương tác giữa các biến giả với nhau.

- **Ví dụ 9:** Hồi quy chi tiêu hàng năm về quần áo (Y) phụ thuộc vào thu nhập (X), giới tính (với hai phạm trù nam và nữ) và thành phần lao động (với hai phạm trù công chức và công nhân).

Mô hình hồi quy: $Y_i = \beta_1 + \beta_2 D_{2i} + \beta_3 D_{3i} + \beta_4 X_i + U_i$

$$D_{2i} = \begin{cases} 1 & \text{nếu đối tượng } i \text{ là nam} \\ 0 & \text{nếu đối tượng } i \text{ là nữ} \end{cases}$$

$$D_{3i} = \begin{cases} 1 & \text{nếu đối tượng } i \text{ là công chức} \\ 0 & \text{nếu đối tượng } i \text{ là công nhân} \end{cases}$$

4.2. Hồi quy với biến giả

- Để phân tích sự ảnh hưởng tương tác giữa các biến giả, ta hồi quy mô hình:

$$Y_i = \beta_1 + \beta_2 D_{2i} + \beta_3 D_{3i} + \beta_4 (D_{2i} * D_{3i}) + \beta_5 X_i + U_i$$

- Để xem có ảnh hưởng tương tác giữa hai biến giả hay không ta kiểm định cặp giả thuyết:

$$\begin{cases} H_0 : \beta_4 = 0 \\ H_1 : \beta_4 \neq 0 \end{cases}$$

- Nếu chấp nhận giả thiết H_1 thì có nghĩa là có sự khác nhau về chi tiêu cho quần áo giữa “nam công chức” và các đối tượng khác (nữ công chức, nữ công nhân, nam công nhân).

4.3. Ứng dụng của biến giả

4.3.1. Sử dụng biến giả để phân tích biến động mùa vụ

Trong kinh tế có nhiều chỉ tiêu mà sự thay đổi của nó mang tính chất mùa vụ.

- Mùa vụ là biến định tính nên để tách biệt ảnh hưởng của yếu tố mùa vụ ta sử dụng kỹ thuật biến giả.
- Nếu yếu tố mùa vụ được xem xét theo tháng (12 tháng/năm) ta sử dụng 11 biến giả.
- Nếu yếu tố mùa vụ được xem xét theo quý (4 quý/năm) ta sử dụng 3 biến giả.

4.3. Ứng dụng của biến giả

- **Ví dụ 10:** hồi quy tiêu dùng về quần áo, dụng cụ gia đình của hộ gia đình (Y) phụ thuộc vào thu nhập (X) và yếu tố mùa vụ (các quý).

$$D_{2i} = \begin{cases} 1 & \text{Nếu quan sát nằm ở quý 2} \\ 0 & \text{Nếu quan sát nằm ở quý khác} \end{cases}$$

$$D_{3i} = \begin{cases} 1 & \text{Nếu quan sát nằm ở quý 3} \\ 0 & \text{Nếu quan sát nằm ở quý khác} \end{cases}$$

$$D_{4i} = \begin{cases} 1 & \text{Nếu quan sát nằm ở quý 4} \\ 0 & \text{Nếu quan sát nằm ở quý khác} \end{cases}$$

- Mô hình hồi quy tổng thể:

$$Y_i = \beta_1 + \beta_2 D_{2i} + \beta_3 D_{3i} + \beta_4 D_{4i} + \beta_5 X_i + U_i \quad (1)$$

4.3. Ứng dụng của biến giả

- Phân tích ảnh hưởng tương tác giữa yếu tố mùa vụ và thu nhập đến chi tiêu ta sử dụng mô hình tổng quát:

$$Y_i = \beta_1 + \beta_2 D_{2i} + \beta_3 D_{3i} + \beta_4 D_{4i} + \beta_5 X_i \\ + \beta_6 (D_{2i} * X_i) + \beta_7 (D_{3i} * X_i) + \beta_8 (D_{4i} * X_i) + U_i \quad (2)$$

- Việc thực hiện các kiểm định đối với các hệ số của mô hình (1) và (2) sẽ cho ta các biết ảnh hưởng của yếu tố mùa vụ cũng như ảnh hưởng tương tác của yếu tố mùa vụ và thu nhập đến chi tiêu.

4.3. Ứng dụng của biến giả

4.3.2. So sánh hai hồi quy

Giả sử có hai hàm hồi quy:

- Giai đoạn 1: $Y_t = \alpha_1 + \alpha_2 X_t + U_t(*)$ với n_1 quan sát
- Giai đoạn 2: $Y_t = \gamma_1 + \gamma_2 X_t + U_t(**)$ với n_2 quan sát

■ Có 4 trường hợp có thể xảy ra:

- Hai hàm hồi quy trùng nhau ($\alpha_1 = \gamma_1, \alpha_2 = \gamma_2$)
- Hai hàm hồi quy song song ($\alpha_1 \neq \gamma_1, \alpha_2 = \gamma_2$)
- Hai hàm hồi quy có cùng hệ số chặn ($\alpha_1 = \gamma_1, \alpha_2 \neq \gamma_2$)
- Hai hàm hồi quy hoàn toàn khác nhau ($\alpha_1 \neq \gamma_1, \alpha_2 \neq \gamma_2$)

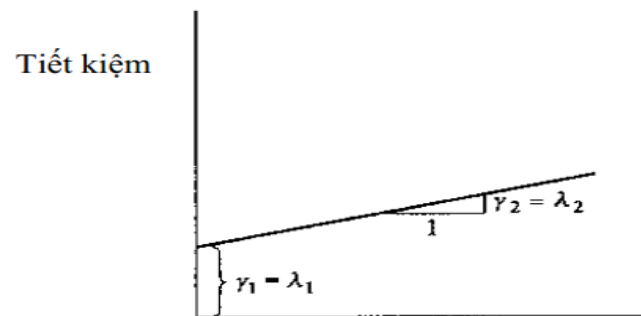
4.3. Ứng dụng của biến giả

- **Ví dụ 11:** Tìm hiểu sự thay đổi về quan hệ giữa tiết kiệm - thu nhập của nước Anh trong hai thời kỳ trước và sau chiến tranh thế giới thứ 2, 1946-1954 (thời kỳ tái thiết) và 1955-1963 (thời kỳ hậu tái thiết).

Bảng 4.2. số liệu về tiết kiệm – thu nhập của Anh

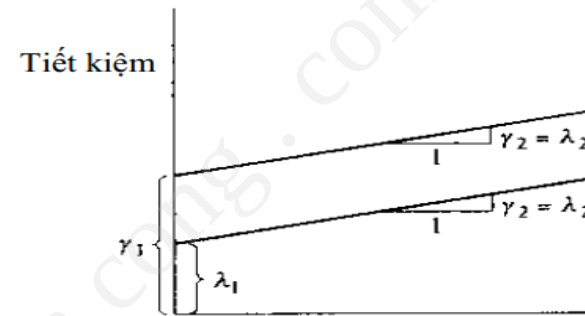
Thời kỳ 1	Tiết kiệm	Thu nhập	Thời kỳ 2	Tiết kiệm	Thu nhập
1946	0.36	8.8	1955	0.59	15.5
1947	0.21	9.4	1956	0.9	16.7
1948	0.08	10	1957	0.95	17.7
1949	0.2	10.6	1958	0.82	18.6
1950	0.1	11	1959	1.04	19.7
1951	0.12	11.9	1960	1.53	21.1
1952	0.41	12.7	1961	1.94	22.8
1953	0.5	13.5	1962	1.75	23.9
1954	0.43	14.3	1963	1.99	25.2

4.3. Ứng dụng của biến giả



(a)

Hồi quy trùng khớp

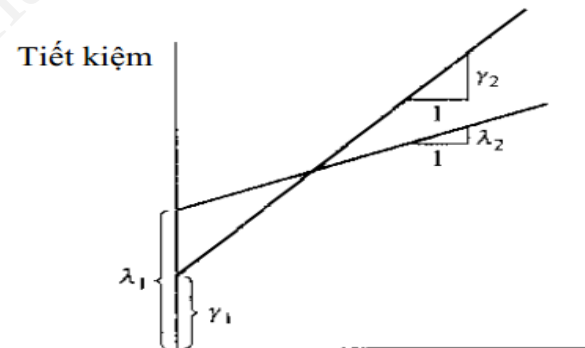
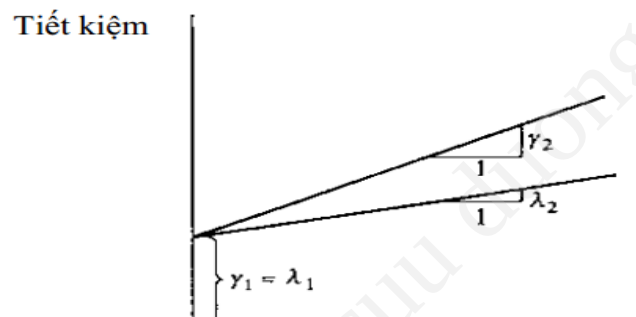


(b)

Hồi quy song song

Thời kỳ tái thiết: $Y_i = \lambda_1 + \lambda_2 X_i + u_{1i}$
 $i = 1, 2, \dots, n_1$

Thời kỳ hậu tái thiết: $Y_i = \gamma_1 + \gamma_2 X_i + u_{2i}$
 $i = 1, 2, \dots, n_2$



Hình 4.2. Đồ thị mối quan hệ giữa tiết kiệm và thu nhập ở các thời kỳ

4.3. Ứng dụng của biến giả

Kiểm định sự đồng nhất giữa hai hồi quy

■ Cách 1: Sử dụng kiểm định Chow:

- Bước 1: Ghép các quan sát của cả hai giai đoạn ta thu được: $n = n_1 + n_2$ quan sát và hồi quy mô hình: $Y_t = \beta_1 + \beta_2 X_t + U_t$ thu được RSS với số bậc tự do $df = (n-k)$.
- Bước 2: Lần lượt hồi quy mô hình tương ứng với hai thời kỳ:
 - Thời kỳ tái thiết thu được RSS_1 với số bậc tự do: $df = (n_1 - k)$
 - Thời kỳ hậu tái thiết thu được RSS_2 với số bậc tự do: $df = (n_2 - k)$

Ký hiệu: $RSS_3 = RSS_1 + RSS_2$ có số bậc tự do: $df = (n_1 - k) + (n_2 - k) = (n - 2k)$.

4.3. Ứng dụng của biến giả

- Bước 3: Kiểm định cặp giả thuyết

$$\begin{cases} H_0 : \text{Cấu trúc của mô hình giữa hai GD là đồng nhất} \\ H_1 : \text{Cấu trúc của mô hình giữa hai GD là không đồng nhất} \end{cases}$$

Tiêu chuẩn kiểm định:

$$F = \frac{(RSS - RSS_3) / k}{RSS_3 / (n - 2k)} \sim F(k, n - 2k)$$

Miền bác bỏ với mức ý nghĩa α cho trước:

$$W_\alpha = \{F, F > F_\alpha(k, n - 2k)\}$$

4.3. Ứng dụng của biến giả

■ Cách 2: Sử dụng biến giả

Đặt:

$$D_t = \begin{cases} 1 & \text{nếu quan sát rơi vào giai đoạn 1} \\ 0 & \text{nếu quan sát rơi vào giai đoạn 2} \end{cases}$$

- Bước 1: Ghép chung các quan sát của cả hai giai đoạn và hồi quy mô hình:

$$Y_t = \beta_1 + \beta_2 D_t + \beta_3 X_t + \beta_4 (D_t * X_t) + U_t$$

Giai đoạn 1: $E(Y_t / X_t, D_t = 1) = (\beta_1 + \beta_2) + (\beta_3 + \beta_4)X_t + U_t$

Giai đoạn 2: $E(Y_t / X_t, D_t = 0) = \beta_1 + \beta_3 X_t + U_t$

4.3. Ứng dụng của biến giả

- Bước 2: Kiểm định các cặp giả thuyết:

$$\begin{cases} H_0 : \beta_2 = \beta_4 = 0 \\ H_1 : \exists \beta_j \neq 0 (j = 2, 4) \end{cases}$$

So sánh cấu trúc của hai hàm hồi quy:

- Nếu cấu trúc của mô hình giữa hai giai đoạn là đồng nhất thì có thể ghép chung số liệu để phân tích.
- Nếu cấu trúc của mô hình giữa hai giai đoạn là không đồng nhất thì phải tách riêng từng tệp số liệu để phân tích.

4.3. Ứng dụng của biến giả

4.3.3. Hồi quy tuyến tính từng khúc

- Nhiều hiện tượng kinh tế có sự thay đổi cấu trúc theo thời gian song vẫn phải đảm bảo tính liên tục theo thời gian của hiện tượng. Khi đó ta sử dụng biến giả để đặc tả cho sự thay đổi cấu trúc này.

- **Ví dụ 12:** Giả sử tiêu dùng của Việt Nam trong 2 thời kỳ trước và sau chuyển đổi (1986) là khác nhau.

Ký hiệu năm chuyển đổi cơ cấu kinh tế là t_0 .

$$D_t = \begin{cases} 1 & \text{nếu } t > t_0 \\ 0 & \text{nếu } t \leq t_0 \end{cases}$$

4.3. Ứng dụng của biến giả

- Xây dựng mô hình:

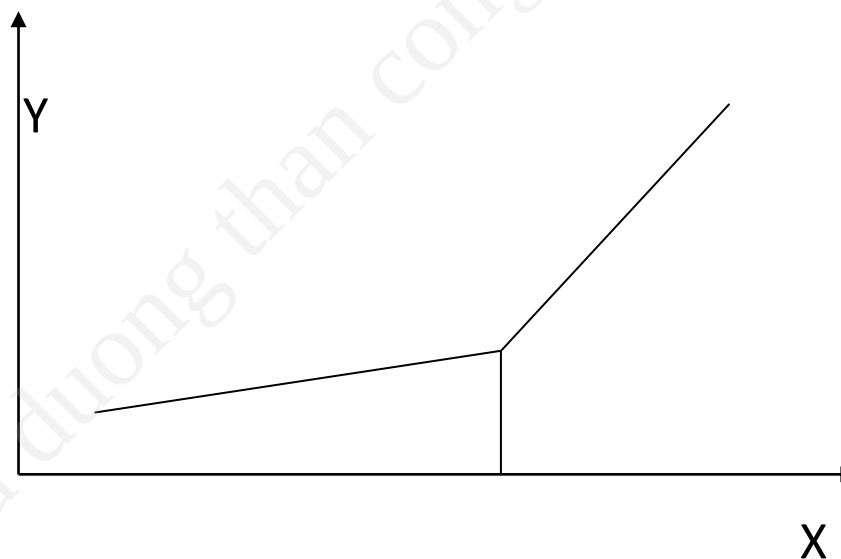
$$Y_t = \beta_1 + \beta_2 X_t + \beta_3 (X_t - X_{t_0}) D_t + U_t$$

Trong đó:

Y_t : Tiêu dùng

X_t : Thu nhập

X_{t_0} : Thu nhập



Hình 4.3. Đồ thị hàm tiêu dùng

4.3. Ứng dụng của biến giả

- Tiêu dùng TB những năm trước chuyển đổi cơ cấu KT.

$$E(Y_t / X_t, D_t = 0) = \beta_1 + \beta_2 X_t$$

- Tiêu dùng TB những năm sau chuyển đổi cơ cấu kinh tế.

$$E(Y_t / X_t, D_t = 1) = (\beta_1 - \beta_3 X_{t_0}) + (\beta_2 + \beta_3) X_t$$

- Tại thời điểm t_0 ta có:

$$E(Y_{t_0}) = (\beta_1 - \beta_3 X_{t_0}) + (\beta_2 + \beta_3) X_{t_0} = \beta_1 + \beta_2 X_{t_0}$$

- Để xem mô hình có thay đổi cấu trúc hay không ta kiểm định cặp giả thiết:

$$\begin{cases} H_0 : \beta_3 = 0 \\ H_1 : \beta_3 \neq 0 \end{cases}$$