Chương 5

HỒI QUY HAI BIẾN: ƯỚC LƯỢNG KHOẢNG VÀ KIỂM ĐỊNH GIẢ THIẾT

Hãy cẩn thận khi kiểm định quá nhiều giả thiết; càng uốn nắn số liệu thì chúng càng dễ cho kết quả, nhưng kết quả thu được bằng cách ép buộc là điều không thể chấp nhận trong khoa học.[1]

Như đã đề cập trong Chương 4, ước lượng và kiểm định giả thiết là hai chuyên ngành lớn của thống kê cổ điển. Lý thuyết ước lượng bao gồm hai phần: ước lượng điểm và ước lượng khoảng. Chúng ta đã thảo luận về ước lượng điểm một cách kỹ lưỡng trong hai chương trước, khi trình bày các phương pháp OLS và ML của ước lượng điểm. Trong chương này, trước hết chúng ta xem xét ước lượng khoảng và sau đó chuyển sang nội dung kiểm định giả thiết, một chủ đề liên quan mật thiết tới ước lượng khoảng.

5.1 CÁC ĐIỀU KIỆN THỐNG KÊ TIÊN QUYẾT

Trước khi minh họa các cơ chế thực sự để thiết lập khoảng tin cậy và kiểm định các giả thiết thống kê, người được đọc xem là đã quen thuộc với các khái niệm cơ bản về xác suất và thống kê. Mặc dù không phải là thay thế cho một khóa học cơ bản về thống kê, Phụ lục A cung cấp các nội dung then chốt của thống kê mà người đọc phải thấu hiểu hoàn toàn. Các khái niệm then chốt như **xác suất, phân phối xác suất**, **sai lầm Loại I và Loại II**, **mức ý nghĩa**, **năng lực của kiểm định thống kê**, và **khoảng tin cậy** rất quan trọng để hiểu các lý thuyết trình bày trong chương này và các chương sau.

5.2 ƯỚC LƯỢNG KHOẢNG: MỘT SỐ KHÁI NIỆM CƠ BẢN

Để làm rõ khái niệm, ta phân tích ví dụ giả thiết về tiêu dùng - thu nhập trong Chương 3. Phương trình (3.6.2) cho thấy xu hướng tiêu dùng biên tế ước lượng (MPC) - β_2 là 0,5091. Đó là một ước lượng đơn (ước lượng điểm) của biến MPC - β_2 của tổng thể chưa biết. Ước lượng này có độ tin như thế nào? Như đã lưu ý trong Chương 3, do các dao động của việc lấy mẫu, một ước lượng đơn có nhiều khả năng khác với giá trị đúng, mặc dù trong việc lấy mẫu lặp lại, giá trị trung bình của nó sẽ bằng với giá trị đúng. (*Lưu ý*: $E(\hat{\beta}_2) = \beta_2$). Trong thống kê, độ tin cậy của một ước lượng điểm được đo bằng sai số chuẩn của nó. Do vậy, thay vì chỉ dựa vào ước lượng điểm, ta có thể xây dựng một khoảng xung quanh giá trị ước lượng điểm, ví dụ trong phạm vi hai hay ba lần sai số chuẩn ở hai phía của giá trị ước lượng điểm, để xác suất mà giá trị đúng của tham số nằm trong khoảng này là, ví dụ, 95%. Đó là sơ bộ ý tưởng đằng sau **ước lượng khoảng**.

Để cụ thể hơn, giả thiết rằng ta muốn tìm xem $(\hat{\beta}_2)$ "gần" với β_2 như thế nào. Để thực hiện mục đích này, ta tìm hai số dương δ và α, số thứ hai nằm trong khoảng từ 0 đến 1, để xác suất mà khoảng ngẫu nhiên ($\hat{\beta}_2 - \delta$, $\hat{\beta}_2 + \delta$) chứa giá trị đúng của β_2 là $1 - \alpha$. Về công thức ta có:

[1] Stephen M. Stigler, "Testing Hypothesis or Fitting Models? Another Look at Mass Extinctions" (Kiểm định giả thiết hay các mô hình thích hợp: một cách nhìn nữa về sự tuyệt chủng), trong *Neutral Models in Biology* (Các mô hình trung lập trong sinh học), Matthew H. Nitecki & Antoni Hoffman hiệu đính, Oxford University Press, Oxford, 1987, trang 148.

$$\Pr(\hat{\beta}_2 - \delta \le \beta_2 \le \hat{\beta}_2 + \delta) = 1 - \alpha \qquad (5.2.1)$$

Khoảng này, nếu tồn tại, được gọi là **khoảng tin cậy**; $1 - \alpha$ được gọi là **hệ số tin cậy**; và α $(0 < \alpha < 1)$ được gọi là **mức ý nghĩa**.[2] Các điểm đầu và cuối của khoảng tin cậy được gọi là các **giới hạn tin cậy** (cũng được gọi là giá trị tới hạn - *critical* value), $\hat{\beta}_2 - \delta$ được gọi là giới hạn **tin cậy dưới** và $\hat{\beta}_2 + \delta$ là giới hạn **tin cậy trên**. Lưu ý rằng α và $1 - \alpha$ thường được biểu diễn dưới dạng phần trăm, 100α và $100(1 - \alpha)$ phần trăm.

Phương trình (5.2.1) cho thấy một ước lượng khoảng, trái với ước lượng điểm, là một khoảng được thiết lập để nó có xác suất chứa giá trị đúng của tham số trong khoảng giới hạn của nó là $1 - \alpha$. Ví dụ, nếu $\alpha = 0{,}05$, hay 5%, (5.2.1) sẽ được phát biểu là: Xác suất mà khoảng (ngẫu nhiên) chỉ ra ở trên chứa giá trị đúng của β_2 là 0,95 hay 95%. Như vậy, ước lượng khoảng cho biết một khoảng các giá trị mà trong đó có thể có giá trị đúng của β_2.

Người đọc cần phải biết các khía cạnh sau đây về ước lượng khoảng:

1. Phương trình (5.2.1) không nói rằng xác suất mà β_2 nằm giữa các giới hạn là $1 - \alpha$. Do β_2, mặc dù chưa biết, được giả thiết là một số cố định, nó có thể nằm ở trong hay ngoài khoảng. Điều mà (5.2.1) diễn đạt là bằng cách sử dụng phương pháp trình bày trong chương này, xác suất của việc xây dựng một khoảng chứa β_2 là $1 - \alpha$.

2. Khoảng (5.2.1) là một **khoảng ngẫu nhiên**, tức là nó thay đổi theo cách chọn mẫu do nó được dựa vào $\hat{\beta}_2$, vốn là một giá trị ngẫu nhiên. (Tại sao?).

3. Do khoảng tin cậy mang tính ngẫu nhiên, các phát biểu về xác suất gắn với nó phải được hiểu theo nghĩa dài hạn, tức là việc lấy mẫu lặp lại. Cụ thể hơn, (5.2.1) mang ý nghĩa là: nếu trong việc lấy mẫu lặp lại, các khoảng tin cậy giống như nó được thiết lập vô số lần trên cơ sở xác suất $1 - \alpha$, thì trong thời gian dài hạn, tính trung bình, có $1 - \alpha$ lần trong tổng số các trường hợp những khoảng này sẽ chứa giá trị đúng của tham số.

4. Như đã nêu ở ý thứ 2, khoảng (5.2.1) là ngẫu nhiên khi $\hat{\beta}_2$ không biết. Nhưng khi ta có một mẫu cụ thể và khi ta tìm được giá trị số học cụ thể của $\hat{\beta}_2$ thì khoảng (5.2.1) không còn ngẫu nhiên nữa; nó được cố định. Trong trường hợp này, ta **không thể** đưa ra phát biểu thống kê (5.2.1); tức là ta không thể nói rằng xác suất mà một khoảng cố định cụ thể chứa giá trị đúng của β_2 là $1 - \alpha$. Trong trường hợp này, β_2 hoặc nằm trong khoảng cố định hay nằm ngoài nó. Do vậy, xác suất là 1 hoặc 0. Như thế, trong ví dụ giả thiết về tiêu dùng - thu nhập, nếu khoảng tin cậy 95% tính được là $(0{,}4268 \le \beta_2 \le 0{,}5941)$, [được giải một cách ngắn gọn trong (5.3.9)}, ta không thể nói rằng xác suất mà khoảng này chứa giá trị đúng của β_2 là 95%. Xác suất đó là 1 hoặc 0.

Các khoảng tin cậy được xây dựng như thế nào? Từ thảo luận ở trên ta có thể đoán rằng nếu việc **lấy mẫu hay phân phối xác suất** của các ước lượng được biết trước, ta có thể đưa ra các phát biểu về khoảng tin cậy như (5.2.1). Trong Chương 4 ta đã thấy với giả thiết phân phối chuẩn của yếu tố nhiễu (hay ngẫu nhiên) u_i, bản thân các ước lượng OLS của $\hat{\beta}_1$ và $\hat{\beta}_2$ có phân phối chuẩn và ước lượng OLS của $\hat{\sigma}^2$ có liên quan phân phối χ^2 (phân phối Chi-bình phương). Từ đó cho thấy công việc thiết lập các khoảng tin cậy có vẻ là một công việc đơn giản. Và sự thật là nó đơn giản!

[2] Cũng được gọi là **xác suất mắc sai lầm Loại I**. Sai lầm Loại I là bác bỏ giả thiết đúng, trái lại sai lầm Loại II là chấp nhận giả thiết sai. (Nội dung này được thảo luận toàn diện hơn trong Phụ lục A). Ký hiệu α được gọi là **kích thước của kiểm định (thống kê)**.

5.3 CÁC KHOẢNG TIN CẬY CHO CÁC HỆ SỐ HỒI QUY β_1 VÀ β_2

Khoảng tin cậy cho β_2

Mục 4.3 trong Chương 4 đã chỉ ra rằng với giả thiết phân phối chuẩn đối với u_i, các ước lượng OLS của $\hat{\beta}_1$ và $\hat{\beta}_2$ tự chúng có phân phối chuẩn với các giá trị trung bình và phương sai tính được. Do đó, ví dụ ta có biến số

$$Z = \frac{\hat{\beta}_2 - \beta_2}{se(\hat{\beta}_2)}$$

$$= \frac{(\hat{\beta}_2 - \beta_2)\sqrt{\sum x_i^2}}{\sigma} \qquad (5.3.1)$$

như đã trình bày trong (4.3.5) là một biến chuẩn đã được chuẩn hóa. Do vậy, có vẻ như ta có thể sử dụng phân phối chuẩn để thực hiện phát biểu xác suất về β_2 với điều kiện là biết phương sai tổng thể σ^2. Nếu σ^2 được biết trước, một tính chất quan trọng của biến có phân phối chuẩn với giá trị trung bình μ và phương sai σ^2 là diện tích ở dưới đường cong chuẩn trong khoảng $\mu \pm \sigma$ bằng gần đúng 68%, trong khoảng $\mu \pm 2\sigma$ bằng gần đúng 95%, và trong khoảng $\mu \pm 3\sigma$ bằng gần đúng 99,7%.

Nhưng σ^2 ít khi được biết trước, và trong thực tế nó được xác định bởi ước lượng không thiên lệch $\hat{\sigma}^2$. Nếu ta thay thế σ bằng $\hat{\sigma}$, (5.3.1) có thể được viết dưới dạng sau:

$$t = \frac{\hat{\beta}_2 - \beta_2}{se(\hat{\beta}_2)} = \frac{\text{ước lượng} - \text{tham số}}{\text{sai số chuẩn của ước lượng tính được}}$$

$$= \frac{(\hat{\beta}_2 - \beta_2)\sqrt{\sum x_i^2}}{\hat{\sigma}} \qquad (5.3.2)$$

với se($\hat{\beta}_2$) bây giờ biểu thị sai số chuẩn ước lượng được. Có thể chỉ ra rằng (xem Phụ lục 5A, Mục 5A.1) biến *t* định nghĩa ở trên tuân theo phân phối *t* với n – 2 bậc tự do. [Lưu ý sự khác nhau giữa (5.3.1) và 5.3.2)]. Do vậy, thay vì sử dụng phân phối chuẩn, ta có thể sử dụng phân phối *t* để thiết lập một khoảng tin cậy cho $\hat{\beta}_2$ như sau:

$$\Pr(-t_{\alpha/2} \le t \le t_{\alpha/2}) = 1 - \alpha \qquad (5.3.3)$$

với giá trị *t* nằm giữa bất đẳng thức kép này là giá trị *t* tính được từ (5.3.2) và với $t_{\alpha/2}$ là giá trị của biến *t* thu được từ phân phối *t* với mức ý nghĩa $\alpha/2$ và n – 2 bậc tự do; nó thường được gọi là giá trị **tới hạn** của *t* tại mức ý nghĩa $\alpha/2$. Thay (5.3.2) vào (5.3.3) ta có:

$$\Pr\left[-t_{\alpha/2} \le \frac{\hat{\beta}_2 - \beta_2}{se(\hat{\beta}_2)} \le t_{\alpha/2}\right] = 1 - \alpha \qquad (5.3.4)$$

Sắp xếp lại (5.3.4) ta có:

$$\Pr[\hat{\beta}_2 - t_{\alpha/2}\,\text{se}(\hat{\beta}_2) \le \beta_2 \le \hat{\beta}_2 + t_{\alpha/2}\,\text{se}(\hat{\beta}_2)] = 1 - \alpha \qquad (5.3.5)$$ [3]

[3] Một số tác giả thích viết (5.3.5) với số bậc tự do được chỉ rõ như sau:

Phương trình (5.3.5) cho biết **khoảng tin cậy** $100(1-\alpha)\%$ của β_2. Ta có thể viết ngắn gọn như sau:

Khoảng tin cậy $100(1-\alpha)\%$ của β_2:

$$\hat{\beta}_2 \pm t_{\alpha/2}\,\mathrm{se}(\hat{\beta}_2) \qquad (5.3.6)$$

Lập luận một cách tương tự và sử dụng (4.3.1) và (4.3.2), ta có thể viết:

$$\Pr[\hat{\beta}_1 - t_{\alpha/2}\,\mathrm{se}(\hat{\beta}_1) \le \beta_1 \le \hat{\beta}_1 + t_{\alpha/2}\,\mathrm{se}(\hat{\beta}_1)] = 1-\alpha \qquad (5.3.7)$$

hay một cách ngắn gọn hơn,

Khoảng tin cậy $100(1-\alpha)\%$ của β_1:

$$\hat{\beta}_1 \pm t_{\alpha/2}\,\mathrm{se}(\hat{\beta}_1) \qquad (5.3.8)$$

Lưu ý một đặc điểm quan trọng của các khoảng tin cậy trình bày trong (5.3.6) và (5.3.8): Trong cả hai trường hợp *chiều rộng của khoảng tin cậy tỷ lệ thuận với sai số chuẩn của ước lượng*. Tức là, sai số chuẩn càng lớn, thì chiều rộng của khoảng tin cậy càng lớn. Nói một cách khác, sai số chuẩn của ước lượng càng lớn thì sự không chắc chắn trong ước lượng giá trị đúng của tham số chưa biết càng lớn. Vì vậy, sai số chuẩn của một ước lượng thường được mô tả là đại lượng đo sự **chính xác** của ước lượng, nghĩa là mức độ chính xác mà ước lượng tính giá trị đúng của tổng thể.

Trở lại ví dụ tiêu dùng - thu nhập trong Chương 3 (Mục 3.6), ta đã tìm ra $\hat{\beta}_2 = 0{,}509$, $\mathrm{se}(\hat{\beta}_2) = 0{,}0357$, và số bậc tự do = 8. Nếu chúng ta giả thiết $\alpha = 5\%$, tức là hệ số tin cậy là 95%, bảng t cho biết với số bậc tự do là 8, giá trị tới hạn $t_{\alpha/2} = t_{0{,}025} = 2{,}306$. Thay những giá trị này vào (5.3.5), người đọc phải tính được khoảng tin cậy 95% của β_2 là:

$$0{,}4268 \le \beta_2 \le 0{,}5914 \qquad (5.3.9)$$

Hay, sử dụng (5.3.6), khoảng tin cậy là:

$$0{,}5091 \pm 2{,}306(0{,}0357)$$

tức là:

$$0{,}5091 \pm 0{,}0823 \qquad (5.3.10)$$

Sự giải thích về khoảng tin cậy này là: với hệ số tin cậy là 95%, trong thời gian dài hạn, 95 trong số 100 trường hợp các khoảng như (0,4268, 0,5914) sẽ chứa giá trị đúng của β_2. Nhưng, như đã cảnh giác ở phần trên, phải chú ý rằng ta không thể nói rằng xác suất khoảng cụ thể (0,4268, 0,5914) chứa giá trị đúng của β_2 là 95% do khoảng này đã được cố định và không còn ngẫu nhiên nữa; do vậy, β_2 hoặc nằm trong khoảng hoặc không: do vậy, xác suất mà khoảng tin cậy cụ thể chứa giá trị đúng của β_2 là 1 hoặc 0.

Khoảng tin cậy đối với β_1

Tương tự như (5.3.7), người đọc có thể dễ dàng chứng minh được rằng khoảng tin cậy 95% của β_1 trong ví dụ tiêu dùng - thu nhập của chúng ta là

$$\Pr[\hat{\beta}_2 - t_{(n-2),\alpha/2}\,\mathrm{se}(\hat{\beta}_2) \le \beta_2 \le \hat{\beta}_2 + t_{(n-2),\alpha/2}\,\mathrm{se}(\hat{\beta}_2)] = 1-\alpha$$

Nhưng để đơn giản ta sẽ giữ nguyên ký hiệu của mình; ngữ cảnh sẽ làm rõ số bậc tự do thích hợp sử dụng.

$$9{,}6643 \le \beta_1 \le 39{,}2448 \qquad (5.3.11)$$

Hay, sử dụng (5.3.8), ta có

$$24{,}4545 \pm 2{,}306(6{,}4138)$$

tức là

$$24{,}4545 \pm 14{,}7902 \qquad (5.3.12)$$

Cũng như trước, người đọc phải cẩn thận khi giải thích khoảng tin cậy này. Trong thời gian dài hạn, 95 trong số 100 trường hợp như (5.3.11) sẽ chứa giá trị đúng của β_1; xác suất mà một khoảng cố định cá biệt chứa giá trị đúng của β_1 là 1 hoặc 0.

Khoảng tin cậy đồng thời cho β_1 và β_2

Có những trường hợp mà ta cần phải thiết lập một khoảng tin cậy đồng thời cho β_1 và β_2 để với hệ số tin cậy $(1-\alpha)$, ví dụ, 95%, cả β_1 và β_2 cùng nằm đồng thời trong khoảng đó. Do nội dung này cũng có liên quan, người đọc có thể muốn xem các tài liệu tham khảo.[4] (Xem đồng thời Mục 8.4 và Chương 10).

5.4 KHOẢNG TIN CẬY ĐỐI VỚI σ^2

Như đã chỉ ra trong Chương 4, Mục 4.3, với giả thiết về phân phối chuẩn, biến

$$\chi^2 = (n-2)\frac{\hat{\sigma}^2}{\sigma^2} \qquad (5.4.1)$$

tuân theo phân phối χ^2 với n – 2 bậc tự do.[5] Do vậy, ta có thể sử dụng phân phối χ^2 để thiết lập khoảng tin cậy cho σ^2

$$\Pr(\chi^2_{1-\alpha/2} \le \chi^2 \le \chi^2_{\alpha/2}) = 1-\alpha \qquad (5.4.2)$$

với giá trị χ^2 nằm giữa bất đẳng thức kép này được tính theo (5.4.1) và với $\chi^2_{1-\alpha/2}$ và $\chi^2_{\alpha/2}$ là hai giá trị của χ^2 (các giá trị **tới hạn** của χ^2) tính được từ bảng Chi-bình phương với n – 2 bậc tự do sao cho chúng cắt ra $100(\alpha/2)$ phần trăm diện tích đuôi của phân phối χ^2, như minh họa trong Hình 5.1.

Thay thế χ^2 từ (5.4.1) vào (5.4.2) và sắp xếp lại các số hạng, ta có

$$\Pr\left[(n-2)\frac{\hat{\sigma}^2}{\chi^2_{\alpha/2}} \le \sigma^2 \le (n-2)\frac{\hat{\sigma}^2}{\chi^2_{1-\alpha/2}}\right] = 1-\alpha \qquad (5.4.3)$$

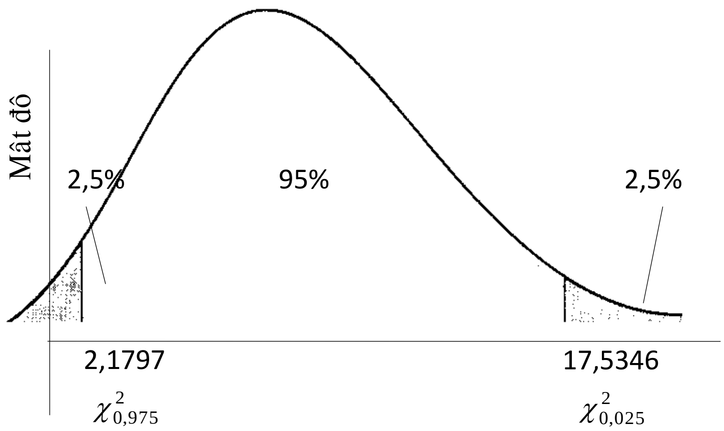
Biểu thức này cho biết khoảng tin cậy $100(1-\alpha)\%$ cho σ^2.

Để minh họa, hãy xem ví dụ sau đây. Trong Chương 3, Mục 3.6, ta tính được $\hat{\sigma}^2 = 42{,}1591$ và số bậc tự do = 8. Nếu α được chọn ở giá trị 5%, bảng Chi-bình phương với số bậc tự do là 8 cho ta các giá trị tới hạn sau: $\chi^2_{0,025} = 17{,}5346$ và $\chi^2_{0,975} = 2{,}1797$. Các giá trị này cho thấy xác suất của một giá trị Chi-bình phương lớn hơn 17,5346 là 2,5% và lớn hơn 2,1797 là 97,5%. Do vậy, khoảng

[4] Xem John Neter, William Wasserman, và Michael H. Kutner, *Applied Linear Regression Models* (Các mô hình hồi quy tuyến tính ứng dụng), Richard D. Irwin, Homewood, Ill., 1983, Chương 5.

[5] Về phần chứng minh, xem Robert V. Hogg & Allen T. Craig, *Introduction to Mathematical Statistics* (Giới thiệu thống kê toán), xuất bản lần thứ 2, Macmillan, New York, 1965, trang 144.

nằm giữa hai giá trị này là khoảng tin cậy 95% của χ^2, như được minh họa bằng đồ thị trong Hình 5.1. (Chú ý tới đặc điểm lệch của phân phối Chi-bình phương).



HÌNH 5.1
Khoảng tin cậy 95% đối với χ^2 (8 bậc tự do)

Thay thế số liệu trong ví dụ của chúng ta vào (5.4.3), người đọc phải tính được khoảng tin cậy 95% của σ^2 như sau:

$$19{,}2347 \le \sigma^2 \le 154{,}7336$$

Sự giải thích về khoảng này là: Nếu ta thiết lập các giới hạn tin cậy 95% đối với σ^2 và nếu ta duy trì một sự tiên nghiệm rằng các giới hạn này sẽ chứa giá trị đúng của σ^2, ta sẽ đúng 95% trong số các trường hợp trong thời gian dài hạn.

5.5 KIỂM ĐỊNH GIẢ THIẾT: CÁC BÌNH LUẬN TỔNG QUÁT

Sau khi đã thảo luận vấn đề ước lượng điểm và ước lượng khoảng, bây giờ ta sẽ xem xét nội dung kiểm định giả thiết. Trong mục này chúng ta thảo luận ngắn gọn một số khía cạnh của chủ đề này; Phụ lục A đưa ra thêm một số chi tiết.

Vấn đề kiểm định giả thiết thống kê có thể được phát biểu đơn giản như sau: *Một quan sát xác định hay kết quả tìm được có tương thích với một giả thiết nêu ra hay không?* Từ "*tương thích*" sử dụng ở đây có nghĩa là "đủ" sát với giá trị được giả thiết để ta không bác bỏ giả thiết phát biểu. Như vậy, nếu một lý thuyết hay kinh nghiệm từ trước làm ta tin rằng hệ số góc đúng β_2 trong ví dụ tiêu dùng - thu nhập là 1đơn vị, thì giá trị quan sát $\hat{\beta}_2$ = 0,5091 tính được từ mẫu trong Bảng 3.2 có phù hợp vơi giả thiết phát biểu không? Nếu có, ta không bác bỏ giả thiết; nếu không, ta có thể bác bỏ nó.

Trong ngôn ngữ thống kê, giả thiết phát biểu được gọi là **giả thiết không** và được ký hiệu là H_0. *Giả thiết không* thường được kiểm định so với một **giả thiết thay thế** ký hiệu H_1(hay còn gọi là **giả thiết đối, giả thiết duy trì**). Ví dụ, giả thiết thay thế H_1 này có thể phát biểu là giá trị đúng của β_2 có thể khác 1. Giả thiết thay thế có thể **đơn giản** hay **phức hợp**.[6] Ví dụ, H_1: β_2 = 1,5 là một giả thiết đơn giản, nhưng H_1: $\beta_2 \neq$ 1,5 là một giả thiết phức hợp.

[6] Một giả thiết thống kê được gọi là **giả thiết đơn giản** nếu nó cụ thể hóa (các) giá trị chính xác của (các) tham số của một hàm mật độ xác suất; nếu ngược lại, giả thiết được gọi là **giả thiết phức hợp**. Ví dụ, trong phân phối chuẩn pdf

Lý thuyết kiểm định giả thiết là xây dựng các quy tắc hay thủ tục để quyết định bác bỏ hay không bác bỏ *giả thiết không*. Có hai cách *tiếp cận bổ sung lẫn nhau* để xây dựng các quy tắc đó, gọi là **khoảng tin cậy** và **kiểm định ý nghĩa**. Cả hai phương pháp này khẳng định rằng biến số (thống kê hay ước lượng) đang xem xét có phân phối xác suất và kiểm định giả thiết là đưa ra các phát biểu hay khẳng định về (các) giá trị hay (các) tham số của phân phối đó, Ví dụ, ta biết rằng với giả thiết về phân phối xác suất chuẩn, thì $\hat{\beta}_2$ có phân phối chuẩn với giá trị trung bình bằng β_2 và phương sai xác định trong (4.3.4). Nếu ta giả thiết là $\beta_2 = 1$, thì ta đang đưa ra một khẳng định về một trong các tham số của phân phối chuẩn, cụ thể là giá trị trung bình. Phần lớn các giả thiết thống kê gặp phải trong cuốn sách sẽ ở vào dạng này – đưa ra các khẳng định về một hay nhiều giá trị của các tham số của một phân phối xác suất giả thiết nào đó như các tham số có phân phối chuẩn, *F, t,* hay χ^2. Các phần sau đây sẽ thảo luận xem làm thế nào để thực hiện được các công việc này.

5.6 KIỂM ĐỊNH GIẢ THIẾT: PHƯƠNG PHÁP KHOẢNG TIN CẬY

Kiểm định hai phía hay hai đuôi

Để minh họa phương pháp khoảng tin cậy, một lần nữa chúng ta trở lại với ví dụ tiêu dùng - thu nhập. Như ta đã biết, xu hướng tiêu dùng biên tế ước lượng được (MPC), $\hat{\beta}_2$, là 0,5091. Giả sử ta mặc định rằng:

$$H_0: \beta_2 = 0{,}3$$
$$H_1: \beta_2 \neq 0{,}3$$

tức là, giá trị đúng của MPC là 0,3 theo *giả thiết không* nhưng nhỏ hơn hay lớn hơn 0,3 theo giả thiết thay thế. *Giả thiết không* là giả thiết đơn giản, trái lại giả thiết thay thế là giả thiết phức hợp; thực tế nó được gọi là **giả thiết hai phía**. Thường thì một giả thiết thay thế có tính chất hai phía phản ánh sự thật là chúng ta không có một nghiên cứu tiên nghiệm hay một kỳ vọng lý thuyết mạnh về hướng đi của giả thiết thay thế xuất phát từ *giả thiết không*.

$\hat{\beta}_2$ quan sát được có *tương thích* với H_0 không? Để trả lời câu hỏi này, hãy tham khảo khoảng tin cậy (5.3.9). Ta biết rằng trong thời gian dài hạn, các khoảng như (0,4268, 0,5914) sẽ chứa giá trị đúng của β_2 với xác suất 95%.

Kết quả là về dài hạn (nghĩa là trong việc lấy mẫu lặp lại), những khoảng như vậy cung cấp một dải hay các giới hạn trong đó giá trị đúng của β_2 có thể nằm trong với một hệ số tin cậy, ví dụ là 95%. Như vậy, khoảng tin cậy cung cấp một tập hợp các các giả thiết H_0 hợp lý. Do đó trong *giả thiết không*, nếu β_2 nằm trong khoảng tin cậy $100(1-\alpha)\%$, ta không bác bỏ *giả thiết không*; nếu nó nằm ngoài khoảng, ta có thể bác bỏ nó.[7] Dải này được minh họa bằng đồ thị trong Hình 5.2.

$(1/\sigma\sqrt{2\pi})\text{mũ}\left\{-\frac{1}{2}\left[(X-\mu)/\sigma\right]^2\right\}$, nếu ta khẳng định rằng H_1: $\mu = 15$ và $\sigma = 2$, nó là một giả thiết đơn giản; nhưng nếu H_1: $\mu = 15$ và $\sigma > 15$, nó là một giả thiết phức hợp, do độ lệch chuẩn không có giá trị cụ thể.

[7] Luôn luôn lưu ý rằng có 100α phần trăm cơ hội mà khoảng tin cậy không chứa β_2 theo H_0 mặc dù giả thiết đúng. Một cách ngắn gọn, có 100α phần trăm cơ hội mắc **sai lầm Loại I**. Như vậy, nếu $\alpha = 0{,}05$, có 5% cơ hội ta có thể bác bỏ *giả thiết không* mặc dù nó đúng.

Các giá trị của β_2 nằm trong khoảng này là hợp lý theo H_0 với độ tin cậy 100(1 − α)%. Do vậy, không bác bỏ H_0 nếu β_2 nằm trong miền này.

$\hat{\beta}_2 - t_{\alpha/2}\text{se}(\hat{\beta}_2)$ $\qquad$ $\hat{\beta}_2 + t_{\alpha/2}\text{se}(\hat{\beta}_2)$

HÌNH 5.2
Khoảng tin cậy 100(1 − α)% của β_2

Quy tắc quyết định: Thiết lập một khoảng tin cậy 100(1 − α) cho β_2. Nếu β_2 theo H_0 nằm trong khoảng tin cậy này, không bác bỏ giả thiết H_0, nhưng nếu β_2 nằm ngoài khoảng này, bác bỏ H_0.

Theo quy tắc này, trong ví dụ giả thiết của chúng ta, H_o: β_2 = 0,3 rõ rằng nằm ngoài khoảng tin cậy 95% cho trong (5.3.9). Do vậy, ta có thể bác bỏ giả thiết rằng giá trị đúng của MPC là 0,3, với độ tin cậy 95%. Nếu giả thiết H_0 đúng, xác suất mà ta có được bằng cách tình cờ một giá trị của MPC như là 0,5091 lớn nhất là 5%, một xác suất nhỏ.

Trong thống kê, khi ta bác bỏ *giả thiết không*, ta nói rằng kết quả của chúng ta có **ý nghĩa thống kê**. Mặc khác, khi ta không bác bỏ *giả thiết không*, ta nói rằng kết quả của chúng ta **không có ý nghĩa thống kê**.

Một số tác giả dùng cụm từ như "rất có ý nghĩa thống kê". Cụm từ này thường có nghĩa là khi bác bỏ *giả thiết không*, xác suất phạm sai lầm Loại I (nghĩa là α) là một số nhỏ, thường là 1%. Nhưng như thảo luận về **giá trị *p*** trong Mục 5.8 sẽ cho thấy, tốt hơn là để cho nhà nghiên cứu quyết định kết quả thống kê là "có ý nghĩa", "khá có ý nghĩa" hay "rất có ý nghĩa".

Kiểm định một phía hay một đuôi

Đôi khi ta có một tiên nghiệm hay kỳ vọng lý thuyết mạnh (hay những kỳ vọng dựa trên một công trình nghiên cứu thực nghiệm trước đó) rằng giả thiết thay thế là một phía hay theo một hướng chứ không phải là hai phía như vừa với thảo luận. Như vậy, trong ví dụ về tiêu dùng - thu nhập, ta có thể viết:

$$H_0: \beta_2 \le 0{,}3 \quad \text{và} \quad H_1: \beta_2 > 0{,}3$$

Có lẽ lý thuyết kinh tế hay công trình nghiên cứu thực nghiệm trước đây cho thấy rằng xu thế tiêu dùng biên tế lớn hơn 0,3. Mặc dù thủ tục để kiểm định giả thiết này có thể được suy ra một cách dễ dàng từ (5.3.5), cách làm thực tế có thể được giải thích một cách tốt hơn theo phương pháp kiểm định ý nghĩa thảo luận ở phần kế tiếp[8]

5.7 KIỂM ĐỊNH GIẢ THIẾT: PHƯƠNG PHÁP KIỂM ĐỊNH Ý NGHĨA

Kiểm định ý nghĩa của các hệ số hồi quy: Kiểm định *t*

Một phương pháp thay thế những bổ sung cho phương pháp khoảng tin cậy để kiểm định các giả thiết thống kê là **phương pháp kiểm định ý nghĩa**. Phương pháp này được phát triển độc lập bởi R. A. Fisher, và hai nhà khoa học Neyman và Pearson.[9] **Nói một cách tổng quát, một kiểm định ý nghĩa là một thủ tục mà các kết quả của mẫu được sử dụng để kiểm chứng tính đúng đắn hay sai lầm của một *giả thiết không*.** Ý tưởng then chốt đằng sau các kiểm định ý nghĩa là một thống kê kiểm định (ước lượng) và phân phối mẫu của thống kê đó theo *giả thiết không*. Quyết định chấp nhận hay bác bỏ H_0 được đưa ra trên cơ sở giá trị của thống kê kiểm định thu được từ số liệu đã có.

Để minh họa, nhớ lại rằng với giả thiết về phân phối chuẩn, biến số

$$t = \frac{\hat{\beta}_2 - \beta_2}{se(\hat{\beta}_2)}$$

$$= \frac{(\hat{\beta}_2 - \beta_2)\sqrt{\sum x_i^2}}{\hat{\sigma}} \qquad (5.3.2)$$

tuân theo phân phối *t* với n – 2 bậc tự do. Nếu giá trị đúng của β_2 được cụ thể hóa theo *giả thiết không*, giá trị *t* trong (5.3.2) có thể hoàn toàn được tính từ mẫu sẵn có, và vì thế mà nó có thể đóng vai trò là một thống kê kiểm định. Do thống kê kiểm định này tuân theo phân phối *t*, ta có thể đưa ra các phát biểu về khoảng tin cậy như sau:

$$\Pr\left[-t_{\alpha/2} \le \frac{\hat{\beta}_2 - \beta_2^*}{se(\hat{\beta}_2)} \le t_{\alpha/2}\right] = 1 - \alpha \qquad (5.7.1)$$

với β_2^* là giá trị của β_2 theo H_0 và với $-t_{\alpha/2}$ và $t_{\alpha/2}$ là các giá trị của *t* (các giá trị **tới hạn** của *t*) tính được từ bảng *t* tại mức ý nghĩa là (α/2) và n – 2 bậc tự do [suy từ (5.3.4)]. Bảng *t* được trình bày trong Phụ lục D.

Sắp xếp lại (5.7.1), ta có

$$\Pr[\beta_2^* - t_{\alpha/2}\mathrm{se}(\hat{\beta}_2) \le \hat{\beta}_2 \le \beta_2^* + t_{\alpha/2}\mathrm{se}(\hat{\beta}_2)] = 1 - \alpha \qquad (5.7.2)$$

Biểu thức này biểu thị khoảng chứa $\hat{\beta}_2$ với xác suất 1 – α, với điều kiện $\beta_2 = \beta_2^*$. Theo ngôn ngữ kiểm định giả thiết, khoảng tin cậy 100(1 – α)% thiết lập trong (5.7.2) được gọi là **miền chấp nhận** (của *giả thiết không*). Và (các) vùng nằm ngoài khoảng tin cậy được gọi là (các) **miền bác bỏ** (của

[8] Nếu bạn muốn sử dụng phương pháp khoảng tin cậy, thiết lập một khoảng tin cậy một phía (100 – α)% cho β_2. Tại sao?

[9] Các chi tiết có thể tìm trong E. L. Lehman, *Testing Statistical Hypotheses* (Kiểm định các giả thiết thống kê), John Wiley & Sons, New York, 1959.

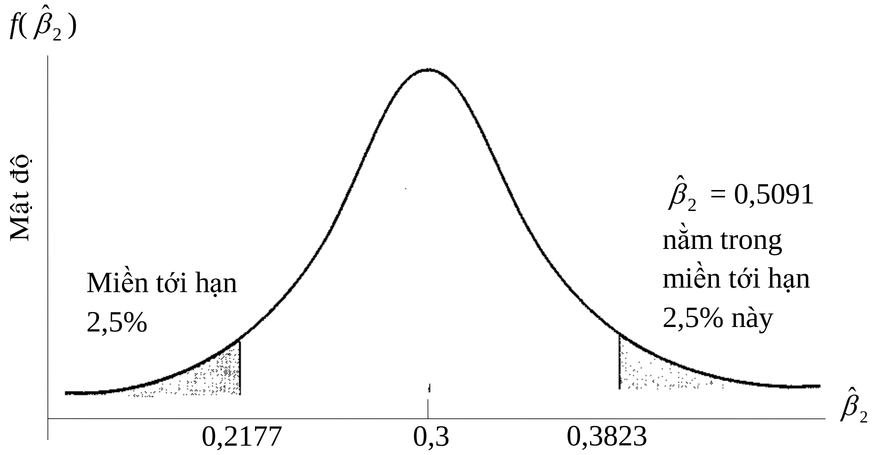
H_0) hay (các) **miền tới hạn**. Như đã lưu ý trước đây, các giới hạn tin cậy, các điểm đầu và cuối của khoảng tin cậy, cũng được gọi là các **giá trị tới hạn**.

Mối liên kết bản chất giữa phương pháp khoảng tin cậy và kiểm định ý nghĩa trong kiểm định giả thiết có thể được nhìn nhận bằng cách so sánh (5.3.5) với (5.7.2). Trong phương pháp khoảng tin cậy, ta thiết lập một dải hay một khoảng chứa giá trị đúng nhưng chưa biết của β_2 với một xác suất nhất định, trái lại trong phương pháp kiểm định ý nghĩa, ta đặt giả thiết một giá trị nào đó của β_2 và xem giá trị tính được $\hat{\beta}_2$ có nằm trong các giới hạn (tin cậy) hợp lý xung quanh giá trị giả thiết hay không.

Một lần nữa hãy trở lại ví dụ về tiêu dùng -thu nhập. Ta biết rằng $\hat{\beta}_2$ = 0,5091, se($\hat{\beta}_2$) = 0,0357, và số bậc tự do = 8. Nếu ta giả sử rằng α = 5%, thì $t_{\alpha/2}$ = 2,306. Nếu ta mặc định H_0: $\beta_2 = \beta_2^*$ = 0,3 và H_1: $\beta_2 \neq$ 0,3, (5.7.2) trở thành

$$\Pr(0{,}2177 \leq \hat{\beta}_2 \leq 0{,}3823) = 0{,}95 \qquad (5.7.3)^{10}$$

như được minh họa bằng đồ thị trong Hình 5.3. Do $\hat{\beta}_2$ quan sát được nằm trong miền tới hạn, ta bác bỏ *giả thiết không* cho rằng giá trị đúng của β_2 = 0,3.



HÌNH 5.3
Khoảng tin cậy 95% đối với $\hat{\beta}_2$ theo giả thiết là β_2 = 0,3

Trên thực tế, không cần phải ước lượng (5.7.2) một cách rõ ràng. Ta có thể tính giá trị *t* nằm ở giữa bất đẳng thức kép (5.7.1) và xem nó có nằm giữa các giá trị tới hạn của *t* hay nằm ngoài chúng. Trong ví dụ của chúng ta:

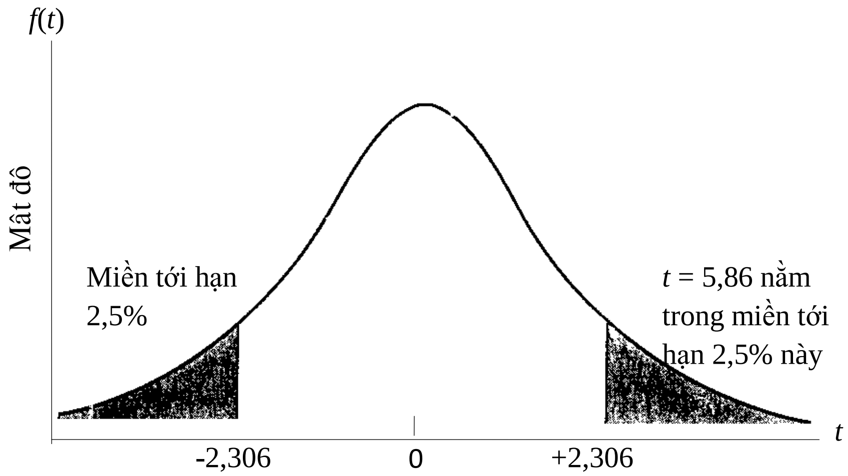
$$t = \frac{0{,}5091 - 0{,}3}{0{,}0357} = 5{,}86 \qquad (5.7.4)$$

Giá trị này rõ ràng nằm trong miền tới hạn của Hình 5.4. Kết luận vẫn như trước, tức là ta bác bỏ H_0.

[10] Mục 5.2, điểm 4 đã phát biểu rằng ta *không thể* nói rằng xác suất mà khoảng cố định (0,4268, 0,5914) chứa giá trị đúng của β_2 là 95%. Nhưng ta có thể đưa ra phát biểu thống kê trình bày trong (5.7.3) do $\hat{\beta}_2$, với tư cách là một ước lượng, là một biến ngẫu nhiên.

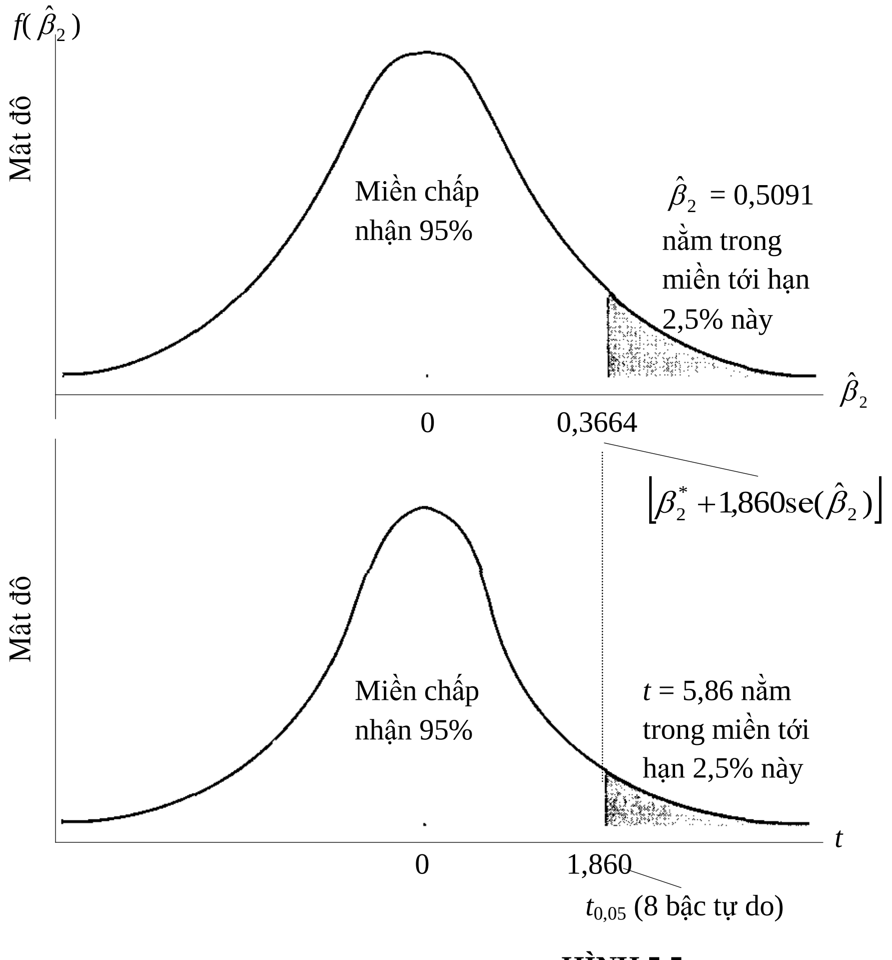
Lưu ý rằng nếu β_2 ước lượng được ($= \hat{\beta}_2$) bằng với giá trị giả thiết β_2, giá trị *t* trong (5.7.4) sẽ bằng 0. Tuy nhiên, do giá trị β_2 ước lượng được khác với giá trị giả thiết của β_2, $|t|$ (tức là, giá trị tuyệt đối của *t*; *Lưu ý*: *t* có thể dương hay âm) sẽ càng lớn. *Do vậy*, một giá trị $|t|$ "lớn" *sẽ là bằng chứng chống lại giả thiết không*. Tất nhiên, ta luôn luôn có thể sử dụng bảng *t* để xác định xem giá trị cá biệt của *t* lớn hay nhỏ; câu trả lời, như ta biết, phụ thuộc vào số bậc tự do cũng như xác suất của sai lầm Loại I mà chúng ta bằng lòng chấp nhận. Nếu bạn xem bảng *t* trong Phụ lục D, bạn sẽ nhận thấy đối với mọi giá trị của bậc tự do, xác suất đạt được các giá trị lớn dần của $|t|$ càng nhỏ dần đi. Như vậy, với 20 bậc tự do, xác suất đạt được giá trị $|t|$ bằng 1,725 hay lớn hơn là 0,10 hay 10%, nhưng với cùng số bậc tự do, xác suất đạt được giá trị $|t|$ bằng 3,552 hay lớn hơn chỉ là 0,002 hay 0,2%.

Do ta sử dụng phân phối *t*, thủ tục kiểm định ở trên được gọi một cách thích hợp là **kiểm định *t*. Trong ngôn ngữ của kiểm định ý nghĩa, một thống kê được *xem* là có ý nghĩa về mặt thống kê nếu giá trị của thống kê kiểm định nằm trong miền tới hạn. Trong trường hợp này, *giả thiết không* bị bác bỏ. Cũng tương tự, một kiểm định được xem là không có ý nghĩa về mặt thống kê nếu giá trị của thống kê kiểm định nằm trong miền chấp nhận.** Trong tình huống này, *giả thiết không* không bị bác bỏ. Trong ví dụ của chúng ta, kiểm định *t* có ý nghĩa và do vậy ta bác bỏ *giả thiết không*.



HÌNH 5.4
Khoảng tin cậy 95% đối với *t* (8 bậc tự do)

Trước khi kết thúc thảo luận về kiểm định giả thiết, Lưu ý rằng thủ tục kiểm định vừa mô tả tóm lược được gọi là **hai phía hay hai đuôi**. Đây là thủ tục kiểm định ý nghĩa trong đó ta xem xét hai phía đuôi của phân phối xác suất, gọi là các miền bác bỏ, và bác bỏ *giả thiết không* nếu nó nằm ở một trong hai phía đuôi. Nhưng điều này xảy ra bởi vì H_1 của ta là giả thiết phức hợp hai phía; $\beta_2 \neq 0,3$, nghĩa là β_2 hoặc lớn hơn hoặc nhỏ hơn 0,3. Nhưng giả sử kinh nghiệm trước đây cho ta thấy rằng MPC được dự kiến là lớn hơn 0,3. Trong trường hợp này ta có: H_0: $\beta_2 \leq 0,3$ và $H_1 > 0,3$. Mặc dù H_1 vẫn là một kiểm định phức hợp, bây giờ H_1 có tính một phía, để kiểm định giả thiết này, ta sử dụng **kiểm định một phía** (phía phải), như minh họa trong Hình 5.5. (Xem đồng thời thảo luận trong Mục 5.6).



HÌNH 5.5
Kiểm định ý nghĩa một phía

Thủ tục kiểm định vẫn nhưng trước trừ việc giới hạn tin cậy phía trên hay giá trị tới hạn bây giờ tương ứng với $t_\alpha = t_{0,05}$, tức là, mức 5%. Như Hình 5.5 minh họa, ta không cần xem xét đuôi phía sau của phân phối t trong trường hợp này. Việc sử dụng kiểm định ý nghĩa một phía hay hai phía phụ thuộc vào một số nghiên cứu tiên nghiệm hay các kinh nghiệm thực nghiệm có trước. (Nhưng chi tiết sẽ được trình bày trong Mục 5.8).

Ta có thể tóm tắt phương pháp kiểm định ý nghĩa t trong kiểm định giả thiết trong Bảng 5.1.

BẢNG 5.1
Kiểm định ý nghĩa *t*: các quy tắc quyết định

Loại giả thiết	***H*₀: Giả thiết không**	***H*₁: Giả thiết thay thế**	**Quy tắc quyết định: Bác bỏ *H*₀ nếu**
Hai phía	$\beta_2 = \beta_2^*$	$\beta_2 \neq \beta_2^*$	$\lvert t \rvert > t_{\alpha/2,\ \mathrm{df}}$
Phía phải	$\beta_2 \leq \beta_2^*$	$\beta_2 > \beta_2^*$	$t > t_{\alpha/2,\ \mathrm{df}}$
Phía trái	$\beta_2 \geq \beta_2^*$	$\beta_2 < \beta_2^*$	$t < -t_{\alpha/2,\ \mathrm{df}}$

Ghi chu: β_2^* là giá trị bằng số giả thiết của β_2.
$\lvert t \rvert$ là giá trị tuyệt đối của *t*.
t_α hay $t_{\alpha/2}$ là giá trị tới hạn của *t* tại mức ý nghĩa α hay $\alpha/2$.
df: bậc tự do, bằng (n – 2) đối với mô hình hai biến, (n – 3) đối với mô hình ba biến, v.v…
Kiểm định giả thiết đối với β_1 có cùng thủ tục.

Kiểm định ý nghĩa của σ^2: Kiểm định χ^2

Với một minh họa khác về phương pháp của luận kiểm định ý nghĩa, xem xét biến số sau:

$$\chi^2 = (n-2)\frac{\hat{\sigma}^2}{\sigma^2} \qquad (5.4.1)$$

BẢNG 5.2
Tóm tắt kiểm định χ^2

***H*₀: Giả thiết không**	***H*₁: Giả thiết thay thế**	**Quy tắc quyết định: Bác bỏ *H*₀ nếu**
$\sigma^2 = \sigma_0^2$	$\sigma^2 > \sigma_0^2$	$\dfrac{\mathrm{df}(\hat{\sigma}^2)}{\sigma_0^2} > \chi^2_{\alpha,df}$
$\sigma^2 \leq \sigma_0^2$	$\sigma^2 < \sigma_0^2$	$\dfrac{\mathrm{df}(\hat{\sigma}^2)}{\sigma_0^2} < \chi^2_{(1-\alpha),df}$
$\sigma^2 \geq \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	$\dfrac{\mathrm{df}(\hat{\sigma}^2)}{\sigma_0^2} > \chi^2_{\alpha/2,df}$ hay $< \chi^2_{(1-\alpha/2),df}$

Ghi chú: σ_0^2 là giá trị của σ^2 theo *giả thiết không*. Chữ nhỏ thứ nhất ở dưới χ^2 ở cột cuối cùng là mức ý nghĩa, và chữ nhỏ thứ hai là bậc tự do. Đây là những giá trị tới hạn của Chi-bình phương. Lưu ý rằng df là (n – 2) đối với mô hình hồi quy hai biến, (n – 3) đối với mô hình hồi quy ba biến, v.v…

Biến này, như đã đề cập trước đây, tuân theo phân phối χ^2 với n – 2 bậc tự do. Trong ví dụ giả thiết, $\hat{\sigma}^2$ = 42,1591 và số bậc tự do = 8. Nếu ta mặc định rằng H_0: σ^2 = 85 so với H_1: $\sigma^2 \neq 85$, phương trình (5.4.1) cho ta thống kê kiểm định đối với H_0. Thay thế các giá trị thích hợp vào (5.4.1), có thể tìm ra được rằng với H_0, χ^2 = 3,97. Nếu ta giả sử α = 5%, các giá trị tới hạn của χ^2 bằng 2,1797 và 17,5346. Do giá trị χ^2 tính được nằm khoảng các giới hạn này, số liệu này hỗ trợ *giả thiết không* và ta không bác bỏ nó. (Xem Hình 5.1). Kiểm định này được gọi là **kiểm định ý nghĩa Chi-bình phương**. Phương pháp kiểm định ý nghĩa χ^2 trong kiểm định giả thiết được tóm tắt trong Bảng 5.2.

5.8 KIỂM ĐỊNH GIẢ THIẾT: MỘT SỐ KHÍA CẠNH THỰC TẾ

Ý nghĩa của việc "chấp nhận" và "bác bỏ" một giả thiết

Nếu trên cơ sở của một kiểm định ý nghĩa, ví dụ kiểm định *t*, ta quyết định chấp nhận *giả thiết không*, tất cả những gì ta phát biểu là trên cơ sở bằng chứng của mẫu ta không có lý do bác bỏ nó; ta không thể nói rằng *giả thiết không* là đúng mà không có nghi ngờ nào. Tại sao? Để trả lời câu hỏi này, hãy quay lại ví dụ của chúng ta về tiêu dùng - thu nhập và giả sử rằng H_0: β_2 (MPC) = 0,50. Bây giờ, giá trị ước lượng của MPC là $\hat{\beta}_2$= 0,5091 với se($\hat{\beta}_2$) = 0,0357. Như vậy, trên cơ sở của kiểm định *t*, ta tìm ra rằng t = (0,5091 − 0,50)/0,0357 = 0,25. *t* không có ý nghĩa tại α = 5%. Do vậy, ta nói "chấp nhận" H_0. Nhưng bây giờ hãy giả sử H_0: β_2 = 0,48. Áp dụng kiểm định, ta có t = (0,5091 - 0,48)/0,0357 = 0,82. Giá trị này cũng không có ý nghĩa thống kê. Và chúng ta cũng nói "chấp nhận" H_0. Giả thiết nào đúng trong hai *giả thiết không* này? Ta không biết. Do vậy, bằng cách chấp nhận *giả thiết không* ta phải luôn luôn nhận thức được rằng một *giả thiết không* nữa cũng có thể hoàn toàn *tương thích* với số liệu. Do vậy, tốt hơn là nên nói rằng ta *có thể* chấp nhận *giả thiết không* chứ không nên nói là chấp nhận nó. Tốt hơn nữa là:

> …cũng như tòa tuyên án là "không phạm tội" chứ không phải là "trong sạch", kết luận của một kiểm định thống kê là "không bác bỏ" chứ không phải là "chấp nhận".[11]

Giả thiết không "zero" và quy tắc kinh nghiệm "2-*t*"

Một *giả thiết không* thường được kiểm định trong nghiên cứu thực nghiệm là H_o: β_2 = 0, tức là hệ số góc bằng không. *Giả thiết không* "zero" này là một loại hình nộm, mục đích là để tìm xem *Y* có quan hệ gì với *X*, biến giải thích, hay không. Nếu bắt đầu từ việc không có quan hệ giữa *Y* và *X* thì việc kiểm định giả thiết như β_2 = 0,3 hay mọi giá trị khác là vô nghĩa.

Giả thiết không này có thể được dễ dàng kiểm định bằng phương pháp khoảng tin cậy hay kiểm định *t* đã được thảo luận trong các phần trên. Nhưng thường thì cách kiểm định chính thức này có thể được làm tắt bằng cách áp dụng quy tắc "2-*t*". Quy tắc này được phát biểu như sau:

> **Quy tắc kinh nghiệm"2-*t*"**. Nếu số bậc tự do lớn hơn hoặc bằng 20 và nếu α, mức ý nghĩa, là 0,05, thì *giả thiết không* β_2 = 0 có thể bị bác bỏ nếu giá trị t [= $\hat{\beta}_2$/se($\hat{\beta}_2$)] tính từ (5.3.2) lớn hơn 2 về giá trị tuyệt đối.

Lý do căn bản của quy tắc này không quá khó chứng minh. Từ (5.7.1) ta biết là sẽ bác bỏ H_0: β_2 = 0 nếu

$$t = \hat{\beta}_2/\text{se}(\hat{\beta}_2) > t_{\alpha/2} \qquad \text{khi } \hat{\beta}_2 > 0$$

hay

$$t = \hat{\beta}_2/\text{se}(\hat{\beta}_2) < -t_{\alpha/2} \qquad \text{khi } \hat{\beta}_2 < 0$$

hay khi

$$|t| = \left|\frac{\hat{\beta}_2}{se(\hat{\beta}_2)}\right| > t_{\alpha/2} \tag{5.8.1}$$

với số bậc tự do phù hợp.

[11] Jan Kmenta, *Elements of Econometrics* (Căn bản về Kinh tế Lượng), Macmillan, New York, 1971, trang 114.

Bây giờ, xem xét bảng *t* trong Phụ lục D, ta thấy rằng với số bậc tự do lớn hơn hoặc bằng 20, giá trị *t* tính được lớn hơn 2 (về trị tuyệt đối), ví dụ như 2,1, sẽ có ý nghĩa thống kê ở mức 5%. Từ đó, ta bác bỏ *giả thiết không*. Do vậy, nếu thấy giá trị tính được của *t* là 2,5 hay 3 với số bậc tự do lớn hơn hoặc bằng 20, ta không cần tra bảng *t* để đánh giá ý nghĩa của hệ số góc tính được. Tất nhiên, người ta luôn luôn có thể tra bảng *t* để tính mức ý nghĩa chính xác, và phải luôn luôn làm vậy nếu số bậc tự do nhỏ hơn 20.

Trước khi chuyển sang phần khác, lưu ý rằng nếu đang kiểm định giả thiết một phía $\beta_2 = 0$ đối lại với $\beta_2 > 0$ hay $\beta_2 < 0$, ta phải bác bỏ *giả thiết không* nếu

$$|t| = \left|\frac{\hat{\beta}_2}{se(\hat{\beta}_2)}\right| > t_\alpha \tag{5.8.2}$$

Nếu ta cố định α ở mức 0,05, từ bảng *t* ta nhận thấy với 20 hay nhiều hơn 20 bậc tự do, một giá trị *t* lớn hơn 1,73 có ý nghĩa thống kê ở mức ý nghĩa 5% (một phía). Do vậy, bất cứ khi nào giá trị *t* lớn hơn 1,8 (về trị tuyệt đối) và số bậc tự do lớn hơn hoặc bằng 20, ta không cần tham khảo bảng *t* để xác định ý nghĩa thống kê của hệ số tính được. Tất nhiên, nếu chọn α ở mức 0,01 hay bất kỳ mức nào khác, ta sẽ phải quyết định về giá trị thích hợp của *t* từ giá trị mốc. Nhưng tới giờ thì người đọc phải có khả năng tự làm được.

Lập *giả thiết không* và giả thiết thay thế[12]

Với các *giả thiết không* và giả thiết thay thế cho trước, việc kiểm định chúng về ý nghĩa thống kê không còn là một điều bí ẩn. Nhưng làm sao có thể thiết lập các giả thiết này? Không hề có một quy tắc bất di bất dịch nào. Thường thì tình huống trong nghiên cứu sẽ gợi ý về tính chất của các *giả thiết không* và giả thiết thay thế. Ví dụ, trong Bài tập 5.16 ta được yêu cầu ước lượng đường thị trường vốn (CML) của lý thuyết đầu tư chứng khoán (portfolio theory), trong đó mặc định rằng $E_i = \beta_1 + \beta_2\sigma_i$ với E = suất sinh lợi kỳ vọng từ cơ cấu đầu tư và σ = độ lệch chuẩn của suất sinh lợi, một thước đo rủi ro. Do suất sinh lợi và rủi ro được dự đoán là có quan hệ đồng biến – rủi ro các cao thì suất sinh lợi càng cao – giả thiết thay thế tự nhiên cho *giả thiết không* ($\beta_2 = 0$) sẽ là $\beta_2 > 0$. Tức là, ta sẽ không xem xét các giá trị β_2 nhỏ hơn 0.

Nhưng hãy xem xét trường hợp mức cầu tiền tệ. Như ta sẽ chỉ ra sau đây, một trong các yếu tố xác định quan trọng của mức cầu tiền tệ là thu nhập. Các nghiên cứu trước đây về hàm cầu tiền tệ đã chỉ ra rằng độ co giãn thu nhập của mức cầu tiền tệ (tỷ số phần trăm thay đổi về mức cầu tiền tệ khi thu nhập thay đổi 1%) thường nằm trong khoảng từ 0,7 đến 1,3. Do vậy, trong một nghiên cứu mới về mức cầu tiền tệ, nếu ta mặc định rằng hệ số co giãn thu nhập β_2 là 1, giả thiết thay thế có thể là $\beta_2 \neq 1$, một giả thiết thay thế hai phía.

Như vậy, ta có thể dựa vào các kỳ vọng lý thuyết hay nghiên cứu kinh nghiệm trước đây hay cả hai để thiết lập các giả thiết. Nhưng mặc dù các giả thiết được thiết lập như thế nào đi nữa thì *điều vô cùng quan trọng là nhà nghiên cứu phải thiết lập những giả thiết trước khi thực hiện điều tra thực nghiệm*. Nếu không, nhà nghiên cứu sẽ phạm phải việc lập luận vòng quanh hay cố ước đoán để cho phù hợp với kết quả. Tức là, nếu thiết lập các giả thiết sau khi xem xét các kết quả thực nghiệm, ta có thể muốn thiết lập giả thiết để biện minh cho kết quả tìm được. Phải tránh cách làm này bằng mọi giá, ít nhất là để tạo sự khách quan khoa học. Hãy lưu ý câu trích dẫn của Stigler ở đầu chương!

[12] Về một thảo luận thú vị về lập giả thiết, xem J. Bradford De Long & Kevin Lang, "Are All Economic Hypotheses False?", *Journal of Political Economy*, (Có đúng là tất cả các giả thiết kinh tế đều sai?, Tạp chí Kinh tế Chính trị), tập 100, số 6, 1992, trang 1257-1272.

Lựa chọn mức ý nghĩa α

Chúng ta phải hiểu rõ từ những thảo luận từ đầu tới đây là việc ta có bác bỏ hay không bác bỏ *giả thiết không* phụ thuộc nhiều vào α, mức ý nghĩa hay *xác suất phạm sai lầm* **Loại I** – xác suất bác bỏ giả thiết đúng. Trong phụ lục A, ta thảo luận toàn diện bản chất của sai lầm Loại I, quan hệ của nó với sai lầm Loại II (xác suất chấp nhận giả thiết sai) và tại sao thống kê cổ điển thường tập trung vào sai lầm Loại I. Nhưng ngay cả như thế, tại sao α lại hay được cố định ở mức 1%, 5%, hay nhiều nhất là 10%? Thực tế là các giả thiết này không có gì là bất khả xâm phạm; mọi giá trị khác cũng có thể được lựa chọn.

Trong một cuốn sách giới thiệu như thế này, không thể thảo luận chi tiết về lý do tại sao lại chọn mức ý nghĩa 1, 5, hay 10%, bởi vì nó sẽ đưa chúng ta tới lĩnh vực ra quyết định thống kê, một lĩnh vực đến từ tự bản thân nó. Tuy nhiên, ta có thể đưa ra một tóm tắt ngắn gọn. Như sẽ thảo luận trong Phụ lục A, với một cỡ mẫu cho trước, nếu ta giảm sai lầm Loại I, sai lầm Loại II tăng lên và ngược lại. Tức là, với cỡ mẫu cho trước, nếu ta giảm xác suất bác bỏ giả thiết đúng, thì đồng thời ta lại tăng xác suất chấp nhận giả thiết sai. Như vậy, có một mối quan hệ được-mất trong hai loại sai lầm này. Bây giờ, cách duy nhất mà chúng ta có thể quyết định về quan hệ được-mất này là tìm chi phí tương đối của hai loại sai lầm. Sau đó,

> Nếu sai lầm bác bỏ *giả thiết không* mà giả thiết đó lại đúng trên thực tế (Sai lầm Loại I) có chi phí cao hơn so với sai lầm không bác bỏ *giả thiết không* khi nó sai trên thực tế (Sai lầm Loại II), việc tạo xác suất loại sai lầm thứ nhất thấp là điều hợp lý. Mặt khác, nếu chi phí của việc phạm Sai lầm Loại I thấp hơn so với chi phí phạm Sai lầm Loại II, sẽ hợp lý nếu tạo xác suất loại sai lầm thứ nhất cao (tức là làm cho xác suất loại sai lầm thứ hai thấp).[13]

Tất nhiên, khó khăn là ở chỗ ta ít khi biết được chi phí của việc phạm hai loại sai lầm. Vì vậy, những nhà kinh tế lượng ứng dụng thường tuân theo cách làm là đặt giá trị của α ở mức 1, hay 5 hay cao nhất là 10% và lựa chọn một thống kê kiểm định mà sẽ làm cho xác suất phạm sai lầm Loại II nhỏ nhất. Bởi vì 1 trừ xác suất phạm sai lầm Loại II được gọi là **năng lực của kiểm định**, cách làm này là để cực đại hóa sức mạnh của kiểm định. (Xem Phụ lục A về phần thảo luận sức mạnh của kiểm định).

Nhưng tất cả vấn đề khó khăn về lựa chọn giá trị thích hợp của α có thể được tránh khỏi nếu ta sử dụng cái gọi là **giá trị *p*** của thống kê kiểm định. Giá trị *p* được thảo luận ở mục kế tiếp.

Mức ý nghĩa chính xác: Giá trị *p*

Như đã lưu ý, "gót chân Asin" của phương pháp cổ điển về kiểm định giả thiết là sự tùy ý trong việc lựa chọn α. Khi ta tính được một thống kê kiểm định (ví dụ thống kê *t*) từ một ví dụ cho trước, tại sao lại không làm theo cách đơn giản là tra bảng thống kê thích hợp để tìm xác suất thực tế của việc đạt được một giá trị của thống kê kiểm định bằng với hay lớn hơn giá trị tính được trong ví dụ? Xác suất này được gọi là **giá trị *p*** (nghĩa là **giá trị xác suất**). Nó cũng được gọi là **mức ý nghĩa quan sát** hay **mức ý nghĩa chính xác** hay **xác suất chính xác phạm sai lầm Loại I**. Nói một cách mang tính kỹ thuật hơn, giá trị *p* được định nghĩa là **mức ý nghĩa thấp nhất mà *giả thiết không* có thể bị bác bỏ**.

Để minh họa, hãy quay lại với ví dụ tiêu dùng - thu nhập. Với *giả thiết không* là giá trị đúng của MPC bằng 0,3, ta có giá trị *t* là 5,86 theo (5.7.4). Giá trị *p* bằng bao nhiêu để đạt được giá trị *t* bằng hay lớn hơn 5,86? Tra bảng *t* trong Phụ lục D, ta thấy với số bậc tự do là 8, xác suất đạt giá trị

[13] Jan Kmenta, *Elements of Econometrics* (Căn bản về Kinh tế Lượng), Macmillan, New York, 1971, trang 126-127.

t như thế phải nhỏ hơn 0,0001 (một phía) hay 0,0002 (hai phía). Bằng cách sử dụng máy tính, có thể chỉ ra rằng xác suất đạt được giá trị *t* bằng 5,86 hay lớn hơn (đối với 8 bậc tự do) vào khoảng 0,000189.[14] Đó là giá trị *p* của thống kê *t*. Mức ý nghĩa quan sát được, hay chính xác của thống kê *t* nhỏ hơn nhiều so với mức ý nghĩa cố định một cách quy ước hay tùy ý, như 1, 5 hay 10%. Trên thực tế, nếu ta sử dụng giá trị *p* vừa tính được và bác bỏ *giả thiết không* cho rằng giá trị đúng của MPC là 0,3, xác suất mà ta phạm sai lầm Loại I chỉ là 0,02%, tức là khoảng 2 trong số 10.000!

Như lưu ý trước đây, nếu số liệu không hỗ trợ *giả thiết không*, $|t|$ tính được theo *giả thiết không* sẽ "lớn" và do vậy giá trị *p* để đạt được *t* như vậy sẽ "nhỏ". Nói một cách khác, với cỡ mẫu cho trước, khi $|t|$ tăng lên, giá trị *p* giảm đi, và do vậy, ta có thể bác bỏ *giả thiết không* với mức tin cậy càng tăng cao.

Đâu là mối quan hệ giữa giá trị *p* và mức ý nghĩa α? Nếu ta tạo thói quen cố định α bằng giá trị *p* của một thống kê kiểm định (ví dụ, thống kê *t*), thì không hề có mâu thuẫn giữa hai giá trị. Nói cách khác, **tốt hơn là từ bỏ cách cố định α một cách tùy ý và đơn giản là chọn giá trị *p* của thống kê kiểm định**. Tốt hơn là để người đọc tự quyết định có bác bỏ *giả thiết không* tại giá trị *p* tính được hay không. Nếu trong một ứng dụng, giá trị *p* của một thống kê kiểm định bằng 0,145 hay 14,5%, và nếu người đọc muốn bác bỏ *giả thiết không* tại mức ý nghĩa (chính xác) này thì cứ việc làm. Không có gì sai nếu chấp nhận xác suất là sẽ sai lầm 14,5% nếu bác bỏ *giả thiết không* trong khi giả thiết đó đúng. Tương tự, như trong ví dụ tiêu dùng - thu nhập của chúng ta, không có gì sai nếu nhà nghiên cứu muốn chọn giá trị *p* vào khoảng 0,02% và không muốn chấp nhận xác suất là phạm sai lầm nhiều hơn 2 trong số 10.000 lần. Nói cho cùng, một số người điều tra có tâm lý thích rủi ro còn số khác lại ghét rủi ro.

Trong phần còn lại của cuốn sách này, nói chung ta sẽ tính giá trị p của một thống kê kiểm định cho trước. Một số người đọc có thể muốn cố định α tại một mức nào đó và bác bỏ *giả thiết không* nếu giá trị *p* nhỏ hơn α. Đó là sự lựa chọn của họ.

Ý nghĩa thống kê so với ý nghĩa thực tế

Hãy quay lại với ví dụ tiêu dùng - thu nhập và lập giả thiết rằng giá trị đúng của MPC là 0,61 (H_0: β_2 = 0,61). Dựa vào kết quả $\hat{\beta}_2$ = 0,5091 trong mẫu, ta có khoảng (0,4268, 0,5914) với 95% độ tin cậy. Do khoảng này không chứa 0,61, ta có thể nói rằng, với 95% độ tin cậy, ước lượng của chúng ta có ý nghĩa thống kê, tức là, kết quả khác đáng kể so với 0,61.

Nhưng đâu là ý nghĩa thực tế hay ý nghĩa lâu dài của kết quả? Tức là, có gì khác khi ta chọn giá trị của MPC là 0,61 chứ không phải là 0,5091? Sự khác biệt 0,1009 giữa hai giá trị MPC có quan trọng trên thực tế không?

Việc trả lời câu hỏi phụ thuộc vào việc ta thực sự làm gì với các ước lượng này. Ví dụ, trong kinh tế vĩ mô ta biết rằng số nhân thu nhập là 1/(1 − MPC). Như vậy, nếu MPC là 0,5091, số nhân là 2,04, nhưng nó sẽ là 2,56 nếu MPC bằng 0,61. Tức là, nếu chính phủ muốn tăng chi tiêu của mình lên 1 USD để đưa nền kinh tế ra khỏi suy thoái, thu nhập sẽ tăng lên 2,04 USD nếu MPC là 0,5091 nhưng sẽ tăng lên 2,56 USD nếu MPC là 0,61. Và như vậy, sự khác biệt có thể rất quan trọng để phục hồi nền kinh tế.

[14] Ta có thể tính giá trị *p* với vài số thập phân bằng cách dùng các bảng thống kê điện tử. Tuy vậy, các bảng thống kê quy ước, do thiếu chỗ, không thể chính xác ở mức đó được. Micro TSP, SHAZAM, ET, và một vài phần mềm thống kê khác có thể tự động cho biết các giá trị *p*.

Điểm Lưu ý trong toàn bộ quá trình thảo luận này là ta không được nhầm lẫn ý nghĩa thống kê với ý nghĩa thực tế, hay kinh tế. Như Goldberger lưu ý:

> Khi một *giả thiết không*, ví dụ $\beta_j = 1$, được cụ thể hóa, người ta thường có ý định cho rằng β_j gần bằng 1, rất gần đến mức mà đối với tất cả các mục đích thực tế, nó có thể được *xem* là *nó bằng 1*. Nhưng 1,1 có "ngang bằng trên thực tế" với 1,0 không là vấn đề kinh tế học, không phải thống kê. Ta không thể giải quyết vấn đề này bằng cách dựa vào một kiểm định thống kê bởi vì thống kê kiểm định $[t = (b_j - 1)/\hat{\sigma}_{bj}]$ tính hệ số ước lượng trong các đơn vị sai số chuẩn. Chúng không phải là các đơn vị có nghĩa để tính hệ số kinh tế $\beta_j - 1$. Tốt hơn là dành thuật ngữ "ý nghĩa" cho khái niệm thống kê, và dùng từ "thực tế" cho khái niệm kinh tế.[15]

Ý tưởng của Goldberger thật sự quan trọng. Khi cỡ mẫu rất lớn, các vấn đề ý nghĩa thống kê trở nên rất ít quan trọng nhưng các vấn đề ý nghĩa kinh tế lại trở nên thiết yếu. Bởi vì với các mẫu rất lớn, hầu hết mọi *giả thiết không* sẽ bị bác bỏ; có thể có các nghiên cứu mà trong đó chỉ quan tâm tới độ lớn của các ước lượng điểm.

Sự lựa chọn giữa phương pháp khoảng tin cậy và kiểm định ý nghĩa trong kiểm định giả thiết thống kê

Trong phần lớn các phân tích kinh tế ứng dụng, *giả thiết không* được thiết lập như là môt hình nộm và mục đích của nghiên cứu thực nghiệm là bác bỏ nó, tức là bác bỏ *giả thiết không*. Như vậy, trong ví dụ tiêu dùng/thu nhập của chúng ta, *giả thiết không* cho rằng MPC, $\beta_2 = 0$ hiển nhiên là ngớ ngẩn, nhưng ta thường sử dụng nó để kịch tính hóa các kết quả thực nghiệm. Rõ ràng là những người biên tập các tạp chí có danh tiếng không lấy gì làm hứng thú khi xuất bản một nghiên cứu thực nghiệm mà lại không bác bỏ *giả thiết không*. Tuy nhiên, kết quả rút ra là MPC khác 0 về mặt thống kê thì lại đáng để đăng tin hơn là tìm ra rằng nó bằng, ví dụ như, 0,7!

Do vậy, J. Bradford Delong và Kevin Lang lập luận rằng tốt hơn là các nhà kinh tế nên

> … tập trung vào trị số của các hệ số và báo cáo về các mức tin cậy chứ không phải các kiểm định ý nghĩa. Nếu tất cả hay gần như tất cả các *giả thiết không* là sai, hoàn toàn có ít giá trị khi ta tập trung vào phân tích xem theo *giả thiết không* thì một ước lượng có thể phân biệt hay không phân biệt với giá trị dự đoán của nó. Thay vào đó, ta muốn làm sáng tỏ những mô hình nào là các phép tính gần đúng tốt. Điều này yêu cầu ta phải biết các khoảng giá trị của thông số mà bị loại trừ bởi các ước lượng thực nghiệm.[16]

Nói tóm lại, các tác giả này thích sử dụng phương pháp khoảng tin cậy hơn so với phương pháp kiểm định ý nghĩa. Người đọc có thể muốn ghi nhớ lời khuyên này.

5.9 PHÂN TÍCH HỒI QUY VÀ PHÂN TÍCH PHƯƠNG SAI

Trong phần này, ta nghiên cứu phân tích hồi quy từ quan điểm phân tích phương sai và giới thiệu cho người đọc một cách nhìn sáng tỏ và mang tính bổ sung về vấn đề suy luận thống kê.

---

[15] Arthur S. Goldberger, *A Course in Econometrics* (Khóa học về Kinh tế lượng), Harvard University Press, Cambridge, Massachusetts, 1991, trang 240. Chú ý b_j là ước lượng OLS của β_j và $\hat{\sigma}_{bj}$ là sai số chuẩn của nó. Về quan điểm chứng thực cho vấn đề này, xem D. N. McCloskey, "The Loss Function Has Been Mislaid: The Rhetoric of Significance Tests" (Hàm số mất đã bị thất lạc: Sự hùng biện của các kiểm định ý nghĩa), *American Economic Review* (Tạp chí Kinh tế Hoa Kỳ), Vol. 75, 1985, trang 201-205.

[16] Xem bài viết của họ trích dẫn trong chú thích 12, trang 1271.

Trong Chương 3, Mục 3.5, ta đã xây dựng đẳng thức sau:

$$\sum y_i^2 = \sum \hat{y}_i^2 + \sum \hat{u}_i^2 = \hat{\beta}_2^2 \sum x_i^2 + \sum \hat{u}_i^2 \qquad (3.5.2)$$

tức là, TSS = ESS + RSS. Đẳng thức này chia tổng bình phương toàn phần thành hai phần: tổng bình phương giải thích được (ESS) và tổng bình phương phần dư (RSS). Nghiên cứu các thành phần này của TSS được gọi là **phân tích phương sai** (ANOVA) từ quan điểm hồi quy.

Liên quan tới mọi tổng bình phương là bậc tự do của nó, tức là số quan sát độc lập mà nó được dựa vào. TSS có n – 1 bậc tự do do ta mất 1 bậc tự do khi tính giá trị trung bình mẫu $\bar{Y}$. RSS có n – 2 bậc tự do. (Tại sao?) (*Lưu ý*: Điều này chỉ đúng với mô hình hồi quy hai biến với sự có mặt của tung độ gốc β_1). ESS có 1 bậc dự do (chỉ đúng cho trường hợp 2 biến). Đó là do ESS = $\hat{\beta}_2^2 \sum x_i^2$ là hàm số của $\hat{\beta}_2$ chỉ khi biết được $\sum x_i^2$.

Hãy sắp xếp các tổng bình phương khác nhau và bậc tự do liên quan của chúng trong Bảng 5.3. Đây là mẫu chuẩn của bảng AOV, đôi khi được gọi là **bảng ANOVA**. Với những công thức trong Bảng 5.3, bây giờ ta xem xét biến số sau:

$$\begin{aligned} F &= \frac{\text{MSS của ESS}}{\text{MSS của RSS}} \\ &= \frac{\hat{\beta}_2^2 \sum x_i^2}{\sum \hat{u}_i^2/(n-2)} \\ &= \frac{\hat{\beta}_2^2 \sum x_i^2}{\hat{\sigma}^2} \end{aligned} \qquad (5.9.1)$$

Nếu giả sử rằng yếu tố nhiễu u_i có phân phối chuẩn và H_0: $\beta_2 = 0$, ta có thể chỉ ra rằng *F* trong (5.9.1) thỏa mãn các điều kiện của Định lý 4.6 (Mục 4.5) và do vậy tuân theo phân phối *F* với 1 và n – 2 bậc tự do. (Xem Phụ lục 5A, Mục 5A.2).

Tỷ số *F* ở trên được dùng để làm gì? Ta có thể chỉ ra rằng[17]

$$E(\hat{\beta}_2^2 \sum x_i^2) = \sigma^2 + \hat{\beta}_2^2 \sum x_i^2 \qquad (5.9.2)$$

và

$$E\frac{\sum \hat{u}_i^2}{n-2} = E(\hat{\sigma}^2) = \sigma^2 \qquad (5.9.3)$$

(Lưu ý rằng β_2 và σ^2 xuất hiện ở vế phải của những phương trình này là các tham số đúng). Do vậy, nếu β_2 bằng 0 trên thực tế, các phương trình (5.9.2) và (5.9.3) cho ta các ước lượng đồng nhất của giá trị đúng của σ^2. Trong tình huống này, biến giả thích *X* không có tác động tuyến tính đối với *Y* và toàn bộ biến thiên của *Y* được giả thích bởi yếu tố nhiễu ngẫu nhiên u_i. Mặt khác, nếu β_2 khác 0, (5.9.2) và (5.9.3) sẽ khác nhau và một phần biến thiên của *Y* sẽ được quy cho *X*. Do vậy, tỷ số *F* trong (5.9.1) cho ta một kiểm định về *giả thiết không* H_0: $\beta_2 = 0$. Do tất cả các số đưa vào phương trình này có thể tính được từ mẫu sẵn có, tỷ số *F* cung cấp một thống kê kiểm định để kiểm định *giả thiết không* cho rằng giá trị đúng của β_2 bằng 0. Tất cả những điều cần phải làm là tính tỷ số *F* và so

[17] Về phần chứng minh, xem K. A. Brownlee, *Statistical Theory and Methodology in Science and Engineering* (Lý thuyết thống kê và phương pháp luận trong khoa học và kỹ thuật), John Wiley & Sons, New York, 1960, trang 278-280.

sánh nó với giá trị *F* tới hạn tính được từ các bảng *F* tại mức ý nghĩa đã chọn, hay thu thập **giá trị *p*** của thống kê *F* đã tính được.

BẢNG 5.3
Bảng ANOVA cho mô hình hồi quy hai biến

Nguồn biến thiên	SS*	Bậc tự do (df)	MSS†
Do hồi quy (ESS)	$\sum \hat{y}_i^2 = \hat{\beta}_2^2 \sum x_i^2$	1	$\hat{\beta}_2^2 \sum x_i^2$
Do phần dư (RSS)	$\sum \hat{u}_i^2$	n – 2	$\dfrac{\sum \hat{u}_i^2}{n-2} = \hat{\sigma}^2$
TSS	$\sum y_i^2$	n – 1	

* SS là tổng bình phương
† Tổng trung bình các bình phương, tính được bằng cách chia SS cho số bậc tự do của nó.

Để minh họa, hãy tiếp tục với ví dụ về tiêu dùng - thu nhập. Bảng ANOVA trong ví dụ này được trình bày trong Bảng 5.4. Giá trị *F* tính được là 202,87. Giá trị *p* của thống kê *F* tương ứng với 1 và 8 bậc tự do không thể tính được từ bảng *F* trong Phụ lục D, nhưng bằng cách sử dụng các bảng thống kê điện tử, ta có thể chỉ ra rằng giá trị *p* là 0,0000001, một xác suất vô cùng nhỏ. Nếu quyết định chọn phương pháp mức ý nghĩa để kiểm định giả thiết và cố định α ở mức 0,01, hay mức 1%, bạn có thể thấy rằng giá trị *F* tính được là 202,87 rõ ràng có ý nghĩa ở mức này. Do vậy, nếu ta bác bỏ *giả thiết không* cho rằng $\beta_2 = 0$, xác suất phạm sai lầm Loại I rất nhỏ. Đối với tất cả các mục đích thực tế, mẫu của chúng ta không thể được chọn từ một tổng thể có giá trị β_2 bằng 0 và ta có thể kết luận với mức tin cậy cao rằng *X*, thu nhập, thật sự có tác động tới *Y*, chi tiêu cho tiêu dùng.

Theo Định lý 4.7 trong Mục 4.5, bình phương giá trị *t* với *k* bậc tự do là giá trị *F* với 1 bậc tự do ở tử số và k bậc tự do ở mẫu số. Trong ví dụ tiêu dùng -thu nhập, nếu giả sử H_0: $\beta_2 = 0$, thì từ (5.3.2) ta có thể dễ dàng chứng minh rằng giá trị ước lượng *t* là 14,24. Giá trị *t* này có 8 bậc tự do. Với cùng *giả thiết không*, giá trị *F* là 202,87 với 1 và 8 bậc tự do. Như vậy, $(14,24)^2$ = giá trị *F*, loại bỏ các sai số do làm tròn.

Vậy, các kiểm định *t* và *F* cho ta hai cách thay thế những bổ sung cho nhau để kiểm định *giả thiết không* là $\beta_2 = 0$. Nếu như vậy thì tại sao lại không chỉ dựa vào kiểm định *t* và bỏ qua kiểm định *F* cùng với phân tích phương sai đi cùng với nó? Đối với mô hình hai biến thì thật sự không cần tới kiểm định *F*. Nhưng khi xem xét chủ đề hồi quy bội, ta sẽ thấy rằng kiểm định *F* có một số ứng dụng thú vị làm cho nó trở thành một phương pháp rất hữu ích và mạnh để kiểm định các giả thiết thống kê.

BẢNG 5.4
Bảng ANOVA cho ví dụ tiêu dùng - thu nhập

Nguồn biến thiên	SS	Bậc tự do	MSS	
Do hồi quy (ESS)	8552,73	1	8552,73	$F = \dfrac{8552{,}73}{42{,}159}$
Do phần dư (RSS)	337,27	8	42,159	= 202,87
TSS	8890,00	9		

5.10 ỨNG DỤNG PHÂN TÍCH HỒI QUY: VẤN ĐỀ DỰ BÁO

Trên cơ sở số liệu mẫu trong Bảng 3.2, ta có hồi quy mẫu sau:

$$\hat{Y}_i = 24{,}4545 + 0{,}5091X_i \tag{3.6.2}$$

với $\hat{Y}_i$ là ước lượng của giá trị đúng $E(Y_i)$ tương ứng với X cho trước. Ta có thể dùng **hồi quy lịch sử** này làm gì? Một cách sử dụng là "dự đoán" hay "dự báo" chi tiêu tiêu dùng trong tương lai Y tương ứng với một mức thu nhập cho trước X. Bây giờ có hai loại dự báo: (1) dự đoán giá trị trung bình có điều kiện của Y tương ứng với một giá trị X cho trước, ví dụ, X_0, tức là điểm trên đường hồi quy tổng thể (xem Hình 2.2), và (2) dự đoán một giá trị cá biệt của Y tướng ứng với X_0. Ta sẽ gọi hai loại dự đoán này là **dự đoán giá trị trung bình** và **dự đoán giá trị cá biệt**.

Dự đoán giá trị trung bình[18]

Để cụ thể hóa, giả sử $X_0 = 100$ và ta muốn dự đoán $E(Y \mid X_0 = 100)$. Bây giờ ta có thể chỉ ra rằng hồi quy lịch sử (3.6.2) cung cấp ước lượng điểm của dự đoán giá trị trung bình này như sau:

$$\begin{aligned}\hat{Y}_0 &= \hat{\beta}_1 + \hat{\beta}_2 X_0 \\ &= 24{,}4545 + 0{,}5091(100) \\ &= 75{,}3645\end{aligned} \tag{5.10.1}$$

với $\hat{Y}_0$ = ước lượng của $E(Y \mid X_0)$. Ta có thể chứng minh rằng giá trị dự đoán điểm này là ước lượng tuyến tính không thiên lệch tốt nhất (BLUE).

Do $\hat{Y}_0$ là một ước lượng, nó có nhiều khả năng khác với giá trị đúng của nó. Sự khác nhau giữa hai giá trị sẽ cho ta một số ý tưởng về sai số dự đoán hay dự báo. Để đánh giá sai số này, ta cần tìm phân phối mẫu của $\hat{Y}_0$. Theo Phụ lục 5A, Mục 5A.3, $\hat{Y}_0$ trong phương trình (5.10.1) có phân phối chuẩn với giá trị trung bình $(\beta_1 + \beta_2 X_0)$ và phương sai tính theo công thức sau:

$$\operatorname{var}(\hat{Y}_0) = \sigma^2\left[\frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum x_i^2}\right] \tag{5.10.2}$$

Bằng cách thay thế σ^2 chưa biết bằng ước lượng không thiên lệch $\hat{\sigma}^2$, ta có

$$t = \frac{\hat{Y}_0 - (\beta_1 + \beta_2 X_0)}{\operatorname{se}(\hat{Y}_0)} \tag{5.10.3}$$

tuân theo phân phối t với n − 2 bậc tự do. Do đó, phân phối t có thể được sử dụng để tính các khoảng tin cậy cho giá trị đúng của $E(Y_0 \mid X_0)$ và kiểm định giả thiết về nó theo cách thông thường. Cụ thể,

$$\Pr[\hat{\beta}_1 + \hat{\beta}_2 X_0 - t_{\alpha/2}\operatorname{se}(\hat{Y}_0) \le \beta_1 + \beta_2 X_0 \le \hat{\beta}_1 + \hat{\beta}_2 X_0 + t_{\alpha/2}\operatorname{se}(\hat{Y}_0)] = 1 - \alpha \tag{5.10.4}$$

với $\operatorname{se}(\hat{Y}_0)$ được tính từ (5.2.10).

---

[18] Về phần chứng minh các phát biểu này, xem Phụ lục 5A, mục 5A.3.

Với số liệu của chúng ta (xem Bảng 3.3),

$$\text{var}(\hat{Y}_0) = 42{,}159\left[\frac{1}{10} + \frac{(100-170)^2}{33.000}\right]$$

$$= 10{,}4759$$

và

$$\text{se}(\hat{Y}_0) = 3{,}2366$$

Do đó, khoảng tin cậy 95% của giá trị đúng của $E(Y \mid X_0) = \beta_1 + \beta_2 X_0$ được tính bởi

$$75{,}3645 - 2{,}306(3{,}2366) \le E(Y \mid X_0 = 100) \le 75{,}3645 + 2{,}306(3{,}2366)$$

tức là,

$$67{,}9010 \le E(Y \mid X_0 = 100) \le 82{,}8381 \qquad (5.10.5)$$

Như vậy, với $X_0 = 100$, trong mẫu lặp lại, 95 trong 100 khoảng giống như (5.10.5) sẽ chứa giá trị trung bình đúng, ước lượng đơn tốt nhất của giá trị trung bình đúng tất nhiên là ước lượng điểm 75,3645.

Nếu ta tính được các khoảng tin cậy 95% như (5.10.5) cho mỗi giá trị X trong Bảng 3.2, ta có cái gọi là **khoảng tin cậy**, hay **dải tin cậy**, cho hàm hồi quy tổng thể. Hàm này được vẽ trong Hình 5.6.

Dự đoán giá trị cá biệt

Nếu sự quan tâm của chúng ta là dự đoán giá trị riêng lẻ của Y, Y_0, tương ứng với một giá trị cho trước của X, ví dụ X_0, thì như trong Phụ lục 5, Mục 5A.3, ước lượng tuyến tích không thiên lệch tốt nhất của Y_0 cũng theo phương trình (5.10.1), nhưng phương sai của nó có giá trị như sau:

$$\text{var}(Y_0 - \hat{Y}_0) = E[Y_0 - \hat{Y}_0]^2 = \sigma^2\left[1 + \frac{1}{n} + \frac{(X_0 - \bar{X}_0)^2}{\sum x_i^2}\right] \qquad (5.10.6)$$

Hơn nữa, Y_0 cũng tuân theo phân phối chuẩn với giá trị trung bình và phương sai tính tương ứng theo (5.10.1) và (5.10.6). Thay $\hat{\sigma}^2$ cho giá trị chưa biết σ^2, suy ra

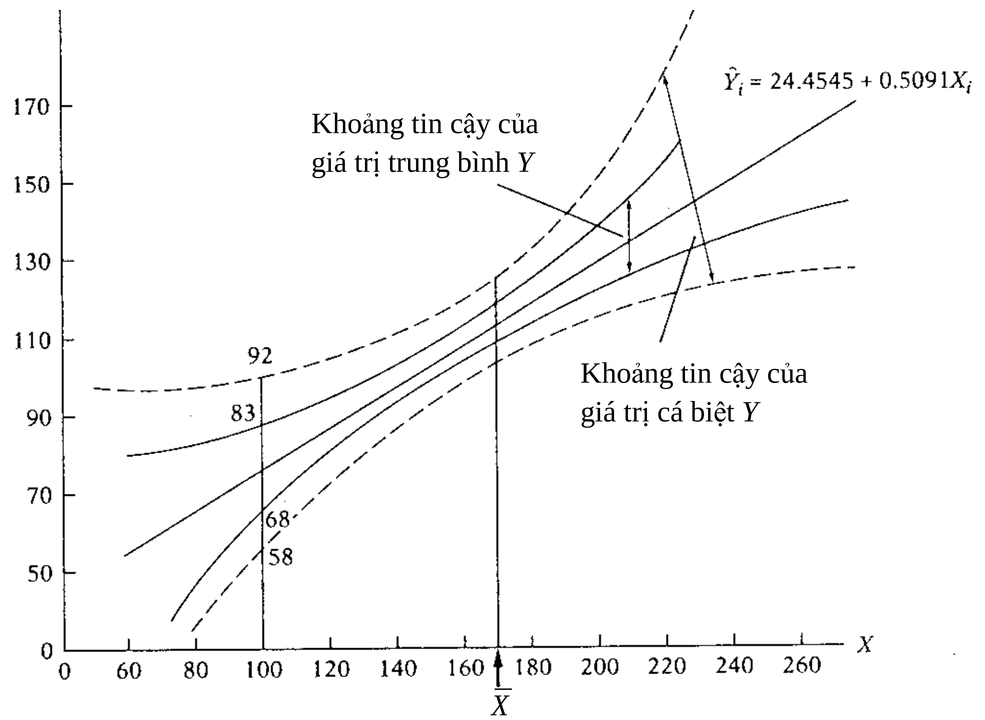
$$t = \frac{Y_0 - \hat{Y}_0}{\text{se}(Y_0 - \hat{Y})}$$

cũng tuân theo phân phối t. Do vậy, phân phối t có thể được dùng để suy luận về giá trị đúng của Y_0. Tiếp tục ví dụ tiêu dùng - thu nhập, ta thấy rằng dự đoán điểm của Y_0 là 75,3645, giống như $\hat{Y}_0$, và phương sai của nó là 52,6349 (người đọc phải chứng minh được phép tính này). Do đó, khoảng tin cậy 95% của Y_0 tương ứng với $X_0 = 100$ là

$$(58{,}6345 \le Y_0 \mid X_0 = 100 \le 92{,}0945) \qquad (5.10.7)$$

So sánh khoảng này với (5.10.5), ta thấy khoảng tin cậy của giá trị riêng lẻ Y_0 rộng hơn khoảng tin cậy của giá trị trung bình Y_0. (Tại sao?) Tính các khoảng tin cậy này giống như (5.10.7) với các giá trị X trong Bảng 3.2, ta có dải tin cậy 95% cho các giá trị riêng lẻ của Y tương ứng với các giá trị của *X*. Dải tin cậy này cùng với dải tin cậy của $\hat{Y}_0$ tương ứng với cùng giá trị *X* được minh họa trong Hình 5.6.

Lưu ý tới đặc điểm quan trọng của các dải tin cậy trong Hình 5.6. Bề rộng của các dải này nhỏ nhất khi $X_0 = \bar{X}$ (Tại sao?) Tuy nhiên, bề rộng lớn lên nhanh chóng khi X_0 tiến xa khỏi $\bar{X}$ (Tại sao?). Sự thay đổi này cho thấy khả năng dự đoán của đường hồi quy mẫu lịch sử giảm mạnh khi X_0 ngày càng xa với $\bar{X}$. **Do vậy, ta phải rất cẩn thận khi "ngoại suy" đường hồi quy lịch sử để dự đoán $E(Y|X_0)$ tương ứng với giá trị cho trước X_0 khác xa với trung bình mẫu $\bar{X}$.**



HÌNH 5.6
Các khoảng (dải) tin cậy của giá trị trung bình của *Y* và giá trị riêng lẻ của *Y*.

5.11 BÁO CÁO CÁC KẾT QUẢ CỦA PHÂN TÍCH HỒI QUY

Có nhiều cách khác nhau để báo cáo các kết quả của phân tích hồi quy, nhưng trong cuốn sách này ta sẽ sử dụng định dạng sau, vận dụng ví dụ tiêu dùng - thu nhập trong Chương 3 như một minh họa:

$$\begin{array}{lllll} \hat{Y}_i = & 24{,}4545 & + \; 0{,}5091X_i & & \\ se = & (6{,}4138) & (0{,}0357) & r^2 = 0{,}9621 & (5.11.1) \\ t = & (3{,}8128) & (14{,}2405) & \text{số bậc tự do (df)} = 8 & \\ p = & (0{,}002571) & (0{,}000000289) & F_{1,8} = 202{,}87 & \end{array}$$

Trong phương trình (5.11.1), các con số trong tập hợp đầu tiên trong ngoặc là các sai số chuẩn ước lượng của các hệ số hồi quy, các con số trong tập hợp thứ hai là các giá trị *t* ước lượng tính từ (5.3.2) theo *giả thiết không* là giá trị tổng thể đúng của mỗi hệ số hồi quy bằng 0 (ví dụ, 3,8128 = 24,4545 ÷ 6,4138), và các số liệu trong tập hợp thứ ba là các giá trị *p* ước lượng. Như vậy, với 8 bậc tự do, xác suất đạt được giá tri *t* là 3,8128 hay lớn hơn là 0,0026 và xác suất đạt được giá trị *t* bằng 14,2405 hay lớn hơn là vào khoảng 0,0000003.

Bằng cách xác định các giá trị *p* của các hệ số *t* ước lượng, ta có thể thấy ngay lập tức mức ý nghĩa chính xác của từng giá trị *t* ước lượng. Như vậy, theo *giả thiết không* là giá trị đúng của tung độ gốc tổng thể bằng 0, xác suất chính xác (nghĩa là giá trị *p*) để đạt được giá trị *t* là 3.8128 hay lớn hơn chỉ vào khoảng 0,0026. Do vậy, nếu ta bác bỏ *giả thiết không* này, xác suất mà ta phạm sai lầm Loại I là vào khoảng 26 trong 10.000, một xác suất rất nhỏ. Đối với tất cả các mục đích thực tế, ta có thể nói rằng giá trị đúng của tung đô gốc tổng thể khác 0. Cũng như vậy, giá trị *p* của hệ số góc ước lượng bằng 0 đối với tất cả các mục đích thực tế. Nếu giá trị đúng của MPC thật sự bằng 0, cơ hội đạt được giá trị MPC là 0,5091 sẽ bằng 0 trên thực tế. Như vậy, ta có thể bác bỏ *giả thiết không* là giá trị đúng của MPC bằng 0.

Trong Định lý 4.7, ta đã chỉ ra mối liên kết cuối cùng giữa thống kê *F* và *t*, tức là, $F_{1,k} = t_k^2$. Theo *giả thiết không* cho rằng giá trị đúng của $\beta_2 = 0$, (5.11.1) cho thấy giá trị *F* là 202,87 (với 1 ở tử số và 8 ở mẫu số) và giá trị *t* vào khoảng 14,24 (8 bậc tự do); như dự kiến, giá trị *F* bằng bình phương giá trị *t*, loại trừ các sai số do làm tròn. Bảng ANOVA đã được thảo luận ở trên.

5.12 ĐÁNH GIÁ CÁC KẾT QUẢ CỦA PHÂN TÍCH HỒI QUY

Trong Hình 1.4 của Phần giới thiệu, ta đã phác họa cơ cấu của việc lập mô hình kinh tế lượng. Ta đã trình bày các kết quả của phân tích hồi quy trong ví dụ tiêu dùng - thu nhập trong (5.11.1). Bây giờ, ta muốn đặt câu hỏi về sự thích hợp của mô hình. Mô hình phù hợp tới đâu? Để trả lời câu hỏi này, ta cần một số tiêu chí.

Thứ nhất, dấu của các hệ số ước lượng có phù hợp với các kỳ vọng lý thuyết hay tiên nghiệm không? Một sự tiên ngiệm là β_2, xu hướng tiêu dùng biên tế (MPC) trong hàm tiêu dùng, phải dương. Trong ví dụ này, β_2 tính được là số dương. Thứ hai, nếu lý thuyết nói rằng mối quan hệ không những chỉ đồng biến mà còn phải có ý nghĩa thống kê thì ví dụ đưa ra có nằm trong trường hợp này không? Như đã thảo luận trong Mục 5.11, MPC không chỉ dương mà còn khác 0 đáng kể về mặt thống kê; giá trị *p* của giá trị *t* ước lượng vô cùng nhỏ. Lập luận cũng đúng cho tung độ gốc. Thứ ba, mô hình hồi quy giải thích biến thiên trong chi tiêu cho tiêu dùng tốt đến đâu? Ta có thể dùng r^2 để trả lời câu hỏi này. Trong ví dụ, r^2 vào khoảng 0.96. Đây là giá trị rất cao khi giá trị cực đại của r^2 là 1.

Như vậy, mô hình mà ta đã lựa chọn để giải thích hành vi chi tiêu tiêu dùng tỏ ra khá tốt. Nhưng trước khi quyết định, ta còn muốn tìm xem mô hình có thỏa mãn các giả thiết về mô hình cổ điển về hồi quy tuyến tính chuẩn (CNLRM) hay không? Ta sẽ không xem xét các giả thiết khác nhau bây giờ bởi vì mô hình rõ ràng quá đơn giản. Nhưng có một giả thiết mà ta cần kiểm tra, đó là quy luật chuẩn của yếu tố nhiễu, u_i. Nhớ lại rằng các kiểm định *t* và *F* sử dụng trước đây yêu cầu rằng sai số tuân theo phân phối chuẩn, Nếu không, thủ tục kiểm định sẽ không có giá trị đối với các mẫu nhỏ, hay mẫu có giới hạn.

Kiểm định quy luật chuẩn

Mặc dù có một số các kiểm định về quy luật chuẩn, ta sẽ chỉ xem xét hai loại: (1) **kiểm định độ phù hợp Chi-bình phương** và (2) **kiểm định Jarque-Bera**. Cả hai kiểm định này đều sử dụng phần dư $\hat{u}_i$ và phân phối xác suất Chi-bình phương.

KIỂM ĐỊNH ĐỘ PHÙ HỢP CHI-BÌNH PHƯƠNG (χ^2).[19] Kiểm định này tiến hành như sau: Trước hết ta chạy hàm hồi quy, tính các phần dư, $\hat{u}_i$, và tính độ lệch chuẩn của $\hat{u}_i$ của mẫu [*Lưu ý*: $\text{var}(\hat{u}_i) = \sum(\hat{u}_i - \bar{\hat{u}})^2 /(n-1) = \sum \hat{u}_i^2 (n-1)$, do $\bar{\hat{u}}$ =0]. Sau đó, ta xếp thứ tự các phần dư và gộp chúng thành các nhóm (trong ví dụ của chúng ta, tất cả có sáu nhóm) tương ứng với số các độ lệch chuẩn khỏi 0. (*Lưu ý*: giá trị trung bình của các phần dư bằng không. Tại sao?) Trong ví dụ của chúng ta, ta có được các số liệu như sau :

Các phần dư quan sát (O_i)	0,0	2,0	3,0	4,0	1,0	0,0	
Các phần dư kỳ vọng (E_i)	0,2	1,4	3,4	3,4	1,4	0,2	
$(O_i - E_i)^2/E_i$	0,2	0,26	0,05	0,10	0,11	0,2	Tổng = 0,92

Lưu ý: $O_i = \hat{u}_i$, với $\hat{u}_i$ là các phần dư OLS.

Dòng "phần dư quan sát" cho biết *phân phối tần suất* của các phần dư đối với các độ lệch chuẩn cụ thể lớn hơn và nhỏ hơn không. Trong ví dụ của chúng ta, không có phần dư nằm ngoài 2 độ lệch chuẩn nhỏ hơn không, 2 phần dư giữa 1 và 2 độ lệch chuẩn nhỏ hơn 0, 3 phần dư giữa 0 và 1 độ lệch chuẩn nhỏ hơn không, 4 phần dư giữa 0 và 1 độ lệch chuẩn lớn hơn 0, 1 phần dư giữa 1 và 2 độ lệch chuẩn lớn hơn 0 và không có phần dư nằm ngoài 2 độ lệch chuẩn lớn hơn 0.

Các số trong dòng phần dư kỳ vọng cho biết phân phối tần suất của các phần dư trên cơ sở của phân phối xác suất giả thiết, trong trường hợp này có dạng chuẩn.[20] Trong hàng thứ ba, ta tính hiệu số giữa các tần suất quan sát và kỳ vọng, bình phương các hiệu số này rồi chia cho các tần suất kỳ vọng và cộng chúng lại. Về mặt đại số, ta có:

$$X^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \qquad (5.12.1)$$

với O_i = tần suất quan sát trong lớp hay khoảng *i* và E_i = tần suất kỳ vọng trong lớp *i* trên cơ sở của phân phối xác suất giả thiết có dạng chuẩn. Bây giờ, nếu hiệu số giữa các tần suất quan sát và kỳ vọng "nhỏ", nó cho thấy yếu tố nhiễu u_i có thể có phân phối xác suất theo giả thiết. Mặt khác, nếu hiệu số giữa các tần suất quan sát và kỳ vọng "lớn", ta có thể bác bỏ *giả thiết không* là các yếu tố nhiễu có phân phối xác suất theo giả thiết. Vì lý do này mà thống kê trong (5.12.1) được gọi là đại lượng đo độ phù hợp bởi vì nó cho biết mức độ mà phân phối xác suất được giả thiết phù hợp với số liệu thực tế (nghĩa là sự phù hợp có tốt không?)

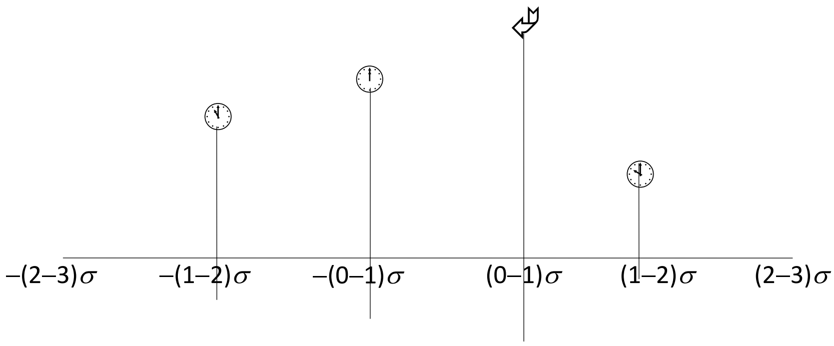
Giá trị X^2 trong (5.12.1) phải có mức độ "lớn" hay "nhỏ" như thế nào trước khi ta quyết định bác bỏ hay không bác bỏ *giả thiết không*? Ta có thể chỉ ra rằng nếu cỡ mẫu tương đối lớn, thống kê X^2 trong (5.12.1) tuân theo *gần đúng phân phối Chi-bình phương (χ^2) với (N – 1) bậc tự do, với N là số lớp hay nhóm*.[21] Một bậc tự do bị mất do điều kiện giới hạn là tổng số các tần suất quan sát và kỳ vọng phải bằng nhau.

[19] Thảo luận sau đây được dựa vào Kenneth J. White & Linda T. M. Bui, *Basic Econometrics: A Computer Handbook Using SHAMZAM* (Kinh tế lượng cơ bản: Sổ tay máy tính sử dụng SHAMZAM) để sử dụng với Gujarati, *Basic Econometrics* (Kinh tế lượng cơ bản), McGraw-Hill, New York, 1988, trang 34. Phần mềm máy tính TSP cũng tuân theo những thủ tục tương tự.

[20] SHAZAM, TSP, ET TM, và một vài phần mềm thống kê có thể đưa ra một phân phối chuẩn phù hợp với một tập hợp số liệu. Các phần mềm này cũng cung cấp kiểm định Chi-bình phương đang được thảo luận một cách ngắn gọn.

[21] Quy tắc chung để tìm các bậc tự do như sau: số bậc tự do = (*N* – 1 – *k*), với *N* là số nhóm và *k* là số tham số ước lượng. Trong trường hợp này, ta đang làm việc với các phần dư $\hat{u}_i$. Nhưng để có những phần dư này, đầu tiên ta phải ước lượng hai đại lượng chưa biết, β_1 và β_2. Vì vậy, ta mất 2 bậc tự do. Bây giờ để làm u_i phù hợp với phân phối chuẩn,

Trở lại ví dụ tiêu dùng - thu nhập, như đã chỉ ra trong bảng ở trên, ta thấy giá trị của X^2 vào khoảng 0,92. Vì chỉ để minh họa nên ta sẽ áp dụng kiểm định Chi-bình phương mặc dù cỡ mẫu khá nhỏ. Ta có sáu nhóm trong ví dụ. Có vẻ như số bậc tự do là (6 − 1) = 5. Nhưng như đã lưu ý trong chú thích 21, ta mất 3 bậc tự do nữa --- 2 do ta phải ước lượng β_1 và β_2 trước khi có thể tính các phần dư $\hat{u}_i$ và 1 do ta đã sử dụng số liệu để ước lượng độ lệch chuẩn của các phần dư. Bây giờ với 2 bậc tự do, **giá trị *p*** để đạt được một giá trị Chi-bình phương bằng hoặc lớn hơn 0,92 là vào khoảng 0,63. Do xác suất này khá cao, sự khác biệt giữa các giá trị quan sát và kỳ vọng của các phần dư không đủ nghiêm trọng để ta bác bỏ giả thiết về quy luật chuẩn.



HÌNH 5.7

Phân phối các phần dư từ ví dụ tiêu dùng - thu nhập, số các độ lệch chuẩn (σ) nhỏ và lớn hơn 0.

Một cách ngẫu nhiên, trước khi áp dụng kiểm định Chi-bình phương vừa mô tả, ta có thể đơn giản vẽ các phần dư quan sát trong bảng ở trên dưới dạng **đồ thị cột** như trong Hình 5.7. Như hình cho thấy, các phần dư quan sát (tính theo đơn vị độ lệch chuẩn khỏi 0) có vẻ gần đúng với phân phối chuẩn. *Thường thì một bức tranh minh họa như thế là một cách tốt để tìm hiểu không chính thức về hình dạng của phân phối xác suất của một biến ngẫu nhiên*.

KIỂM ĐỊNH JARQUE-BERA VỀ QUY LUẬT CHUẨN.[22] Kiểm định JB về quy luật chuẩn là một kiểm định tiệm cận hay kiểm định mẫu lớn. Nó cũng được dựa trên các phần dư OLS. Trước hết, kiểm định này tính **độ lệch** (slewness) và **độ nhọn** (kurtosis) của phân phối xác suất (mô tả trong Phụ lục A) của các phần dư OLS và sử dụng thống kê kiểm định sau:

$$\text{JB} = n\left[\frac{S^2}{6} + \frac{(K-3)^2}{24}\right] \qquad (5.12.2)$$

với *S* là độ lệch và *K* làđộ nhọn.

---

ta phải ước lượng các tham số của phân phối chuẩn, tức là giá trị trung bình và phương sai. Nhưng do giá trị trung bình của $\hat{u}_i$ bằng 0 (tại sao?), ta chỉ phải tính phương sai. Do vậy, ta mất 1 bậc tự do. Từ đó, ta mất *k* = 3 bậc tự do. Với *N* = 6, số bậc tự do là (6 − 1 − 3) = 2. Về cách sử dụng kiểm định Chi-bình phương để tính độ phù hợp, xem mọi cuốn sách giới thiệu về thống kê.

[22] Xem C. M. Jarque & A. K. Bera, "A Test for Normality of Obserations and Regression Residuals" (Một kiểm định quy luật chuẩn của các quan sát và phần dư hồi quy), *International Statistical Review* (Tạp chí Thống kê Quốc tế), số 55, 1987, trang 163-172.

Do đối với một phân phối chuẩn giá trị của độ lệch bằng 0 và giá trị của độ nhọn bằng 3, (*K* – 3) đại diện cho độ nhọn trội trong (5.12.2). Theo *giả thiết không* là các phần dư có phân phối chuẩn, Jarque và Bera đã chỉ ra rằng một cách tiệm cận (nghĩa là trong các mẫu lớn), thống kê JB trong (5.12.2) tuân theo phân phối Chi-bình phương với 2 bậc tự do. Nếu giá trị *p* của thống kê Chi-bình phương tính được trong một ứng dụng có giá trị đủ nhỏ, ta có thể bác bỏ giả thiết là các phần dư có phân phối chuẩn. Nhưng nếu giá trị *p* tương đối lớn, ta không bác bỏ giả thiết về quy luật chuẩn.

Trở lại với ví dụ tiêu dùng - thu nhập, ta tính được (sử dụng phần mềm **SHAZAM**, **TSP**, hay **ET**) giá trị JB là 0,7769. Nếu mẫu tương đối lớn, giá trị *p* để đạt được một giá trị Chi-bình phương như vậy với 2 bậc tự do vào khoảng 0,6781, một xác suất khá lớn. Do vậy, một cách tiệm cận ta không bác bỏ giả thiết về quy luật chuẩn.

Các kiểm định khác về sự phù hợp của mô hình

Nhớ rằng mô hình cổ điển về hồi quy tuyến tính chuẩn (CNLRM) đưa ra nhiều giả thiết chứ không chỉ quy luật chuẩn của số hạng sai số. Khi phát triển lý thuyết kinh tế lượng sâu hơn, ta sẽ xem xét một vài kiểm định khác về sự phù hợp của mô hình. Cho tới khi đó, hãy lưu ý rằng việc lập mô hình hồi quy của chúng ta được dựa vào một vài giả thiết đơn giản hóa, những giả thiết có thể không đúng trong từng trường hợp cụ thể.

5.13 TÓM TẮT VÀ KẾT LUẬN

1. Ước lượng và kiểm định giả thiết là hai nhánh quan trọng của thống kê cổ điển. Sau khi đã thảo luận vấn đề trong Chương 3 và 4, ta phân tích vấn đề kiểm định giả thiết trong chương này.
2. Kiểm định giả thiết trả lời câu hỏi sau: kết quả tìm được có tương thích với một giả thiết đã được phát biểu hay không?
3. Có hai phương pháp bổ sung cho nhau để trả lời câu hỏi ở trên: **khoảng tin cậy** và **kiểm định ý nghĩa**.
4. Đằng sau phương pháp khoảng tin cậy là khái niệm **ước lượng khoảng**. Một ước lượng khoảng là một khoảng hay một dải được thiết lập làm sao để có một xác suất cụ thể chứa trong giới hạn của nó giá trị đúng của một tham số chưa biết. Khoảng này được gọi là **khoảng tin cậy**, thường được phát biểu dưới dạng phần trăm, như 90 hay 95%. Khoảng tin cậy cung cấp một tập hợp các giả thiết hợp lý về giá trị của tham số chưa biết. Nếu giá trị theo *giả thiết không* nằm trong khoảng tin cậy, *giả thiết không* sẽ không bị bác bỏ, trái lại nếu nó nằm ngoài khoảng này, *giả thiết không* có thể bị bác bỏ.
5. Trong thủ tục **kiểm định ý nghĩa**, người ta xây dựng một **thống kê kiểm định** và xem xét phân phối mẫu của nó theo *giả thiết không*. Kiểm định thống kê thường tuân theo một phân phối xác suất xác định như phân phối chuẩn, *t*, *F*, hay Chi-bình phương. Khi một thống kê kiểm định (ví dụ thống kê *t*) được tính từ số liệu đã có, giá trị *p* của nó có thể được tính một cách dễ dàng. Giá trị *p* cho ta xác suất chính xác để đạt được một kiểm định thống kê ước lượng theo *giả thiết không*. Nếu giá trị *p* này nhỏ, ta có thể bác bỏ *giả thiết không*, nhưng nếu giá trị *p* lớn ta có thể không bác bỏ nó. Trong việc lựa chọn giá trị *p* nhà điều tra phải lưu ý tới xác suất phạm **sai lầm Loại I và Loại II**.

6. Trên thực tế, ta phải cẩn thận khi cố định α, xác suất phạm **sai lầm Loại I**, ở những giá trị tùy ý như 1, 5, hay 10%. Tốt hơn là đưa ra **giá trị *p*** của thống kê kiểm định. Đồng thời, ý nghĩa thống kê của một ước lượng không thể được nhầm lẫn với ý nghĩa thực tiễn của nó.
7. Tất nhiên, kiểm định giả thiết mặc định rằng mô hình lựa chọn cho nghiên cứu thực nghiệm là phù hợp trên khía cạnh là nó không vi phạm một hay nhiều giả thiết của mô hình cổ điển về hồi quy tuyến tính chuẩn. Chương này giới thiệu một trong các kiểm định đó, **kiểm định quy luật chuẩn**, để tìm xem đại lượng sai số có tuân theo phân phối chuẩn hay không. Do trong các mẫu nhỏ, hay mẫu giới hạn, các kiểm định *t, F,* và Chi-bình phương yêu cầu giả thiết về quy luật chuẩn, việc kiểm định giả thiết này một cách chính thức là điều quan trọng.
8. Nếu mô hình phù hợp trên thực tế, nó có thể được sử dụng cho các mục đích dự báo. Nhưng trong khi dự báo các giá trị tương lai của biến được hồi quy, ta không được đi quá xa khỏi phạm vi mẫu của các giá trị dùng làm hồi quy. Nếu không, các sai số dự báo có thể tăng lên khá mạnh.

BÀI TẬP

Câu hỏi

5.1. Cho biết và giải thích lý do các phát biểu sau là đúng, sai, hay không chắc chắn. Hãy phát biểu chích xác.

(a) Kiểm định ý nghĩa *t* thảo luận trong chương này yêu cầu rằng phân phối các ước lượng $\hat{\beta}_1$ và $\hat{\beta}_2$ của mẫu tuân theo phân phối chuẩn.
(b) Mặc dù yếu tố nhiễu trong mô hình cổ điển về hồi quy tuyến tính chuẩn (CNRM) không có phân phối chuẩn, các ước lượng OLS vẫn không thiên lệch.
(c) Nếu mô hình hồi quy không có tung độ gốc, các u_i ước lượng (= $\hat{u}_i$) sẽ không có tổng bằng không.
(d) Giá trị *p* và độ lớn của thống kê kiểm định có nghĩa như nhau.
(e) Trong một mô hình hồi quy có tung độ gốc, tổng các phần dư luôn luôn bằng không.
(f) Nếu một *giả thiết không* không bị bác bỏ thì nó đúng.
(g) Giá trị σ^2 càng cao thì phương sai $\hat{\beta}_2$ trong (3.3.1) càng lớn.
(h) Trung bình có điều kiện và không có điều kiện của một biến ngẫu nhiên có nghĩa như nhau.
(i) Trong hàm hồi quy tổng thể hai biến PRF, nếu hệ số góc β_2 bằng không, tung độ gốc được ước lượng bằng giá trị trung bình mẫu $\overline{Y}$
(j) Phương sai có điều kiện, var($Y_i \mid X_i$) = σ^2, và phương sai không điều kiện của Y, var(*Y*) = σ_Y^2 , sẽ như nhau nếu *X* không có tác động tới *Y*.

5.2. Thiết lập bảng ANOVA theo cách trong Bảng 5.4 cho mô hình hồi quy trong (3.7.2) và kiểm định giả thiết cho rằng chi tiêu tiêu dùng cá nhân và tổng sản phẩm quốc dân của nền kinh tế Hoa Kỳ trong giai đoạn 1980-1991 không có liên quan.

5.3. Xem xét các kết quả hồi quy sau đối với nền kinh tế Hoa Kỳ trong giai đoạn 1968-1987 ($\hat{Y}$ = chi tiêu của Hoa Kỳ đối với hàng hóa nhập khẩu và *X* = thu nhập khả dụng cá nhân, cả hai đều tính bằng tỷ đô la, giá cố định 1982):

$$\hat{Y}_t = -261{,}09 \quad + \quad 0{,}2453X_t$$

$$\text{se} = (31{,}327) \qquad (\qquad) \qquad r^2 = 0{,}9388$$

$$t = (\qquad) \qquad (16{,}616) \qquad n = 20$$

(a) Điền số vào các ô trống trong ngoặc.
(b) Bạn giải thích hệ số 0,2453 như thế nào? Và hệ số –261,09?

(c) Có bác bỏ giả thiết cho rằng giá trị đúng của hệ số góc bằng 0 hay không? Sử dụng kiểm định nào? Tại sao? Giá trị *p* của thống kê kiểm định của bạn bằng bao nhiêu?
(d) Thiết lập bảng ANOVA cho ví dụ này và kiểm định giả thiết cho rằng giá trị đúng của hệ số góc bằng 0. Bạn sử dụng kiểm định nào và tại sao?
(e) Các câu trả lời trong (a) và (b) có mẫu thuẫn với nhau không? Nếu không, điều gì giải thích sự điều hòa giữa hai câu trả lời?
(f) Giả sử trong hồi quy vừa nêu, giá trị r^2 không được cho biết. Bạn có thể tính được nó từ các kết quả khác trong hồi quy không?

5.4. Gọi ρ^2 là giá trị đúng của hệ số tương quan tổng thể. Giả sử bạn muốn kiểm định giả thiết là $\rho = 0$. Giải thích bằng lời nói làm thế nào bạn kiểm định giả thiết này. Gợi ý: Sử dụng phương trình (3.5.11). Xem đồng thời bài tập 5.7.

5.5. Cái được gọi là đường đặc tính của phân tích đầu tư hiện đại chỉ đơn giản là đường hồi quy thiết lập từ mô hình sau:

$$r_{it} = \alpha_i + \beta_i r_{mt} + u_t$$

với r_{it} = suất sinh lợi của chứng khoán thứ *i* vào thời gian *t*
r_{mt} = suất sinh lợi trung bình của các chứng khoán trên thị trường vào thời gian *t*
u_t = yếu tố nhiễu ngẫu nhiên

Trong mô hình này β_i được gọi là **hệ số bê ta** của chứng khoán thứ *i,* một đại lượng đo rủi ro thị trường (hay hệ thống) của một chứng khoán.[*]

Dựa vào 240 suất sinh lợi hàng tháng trong thời kỳ 1956-1976, Fogler & Ganapathy tính đường đặc tính sau cho cổ phiếu IBM trong quan hệ với chỉ số chứng khoán thị trường do Đại học Chicago xây dựng:[†]

$$r_{it} = 0{,}7264 + 1{,}0598 r_{mt} \qquad r^2 = 0{,}4710$$
$$\text{se} = (0{,}3001) \quad (0{,}0728) \qquad \text{số bậc tự do} = 238$$
$$F_{1,238} = 211{,}896$$

(a) Một chứng khoán có hệ số bê ta lớn hơn 1 được gọi là chứng khoán dễ biến động hay năng động. IBM có phải là chứng khoán dễ biến động không trong khoảng thời gian ta nghiên cứu?
(b) Hệ số tung độ gốc có khác 0 về ý nghĩa hay không? Nếu có, ý nghĩa thực tế của nó là gì?

5.6. Phương trình (5.3.5) cũng có thể viết dưới dạng

$$\Pr[\hat{\beta}_2 - t_{\alpha/2}\text{se}(\hat{\beta}_2) < \beta_2 < \hat{\beta}_2 + t_{\alpha/2}\text{se}(\hat{\beta}_2)] = 1 - \alpha$$

Tức là, bất đẳng thức yếu (≤) được thay thế bằng bất đẳng thức mạnh (<). Tại sao?

5.7. R. A. Fisher đã tính phân phối mẫu của hệ số tương quan định nghĩa trong (3.5.13). Nếu giả sử rằng các biến *X* và *Y* đồng thời phân phối xác suất theo quy luật chuẩn, tức là nếu chúng thuộc phân phối chuẩn của hai đại đượng ngẫu nhiên (xem Phụ lục 4A, bài tập 4.1), thì theo giả thiết là hệ số tương quan tổng thể ρ bằng 0, có thể chứng minh rằng $t = r\sqrt{n-2}/\sqrt{1-r^2}$ tuân theo phân phối *t* của Student với n – 2 bậc tự do.[**] Chứng minh rằng giá trị *t* đồng nhất với giá trị *t* trong (5.3.2) trong *giả thiết không* cho rằng $\beta_2 = 0$. Sau đó, chứng tỏ rằng với cùng *giả thiết không* $F = t^2$. (Xem Mục 5.9).

[*] Xem Haim Levy & Marshall Sarnat, *Portfolio and Investment Selection: Theory and Pratice* (Cơ cấu chứng khoán và lựa chọn đầu tư: lý thuyết và thực tiễn), Prentice-Hall International, Englewood Cliffs, N.J., 1984, Chương 12.

[†] H. Russell Fogler & Sundaram Ganapathy, *Financial Econometrics* (Kinh tế lượng tài chính), Prentice-Hall, Englewood Cliffs, N.J., 1982, trang 13.

[**] Nếu ρ thật sự bằng 0, Fisher đã chứng minh rằng *r* tuân theo cùng một phân phối *t* với điều kiện một trong hai biến *X* hay *Y* có phân phối chuẩn. Nhưng nếu ρ khác 0, cả hai biến phải có phân phối chuẩn. Xem R. L. Anderson & T. A.

Bài tập

5.8. Tham khảo hàm cầu đối với cà phê trong phương trình (3.7.1).

(a) Thiết lập các khoảng tin cậy 95% *riêng lẻ* cho β_1, β_2 và σ^2.
(b) Sử dụng phương pháp khoảng tin cậy để kiểm định giả thiết cho rằng giá cà phê không có tác động gì tới mức tiêu dùng cà phê.
(c) Làm lại câu (b), sử dụng phương pháp kiểm định ý nghĩa. Bạn sử dụng kiểm định nào và tại sao? Lấy $\alpha = 5\%$.
(d) Giá trị p của thống kê kiểm định tính trong câu (c) bằng bao nhiêu? Nếu giá trị p này nhỏ hơn α, bạn có kết luận gì?
(e) Thiết lập bảng ANOVA cho bài tập này và kiểm định giả thiết $\beta_2 = 0$. Có mâu thuẫn giữa câu trả lời này với câu (b) không?
(f) Thay cho việc kiểm định *giả thiết không* $\beta_2 = 0$, bạn có thể kiểm định giả thiết cho rằng giá trị đúng của hệ số tất định (coefficient of determination) bằng 0 được không?
(g) Giả sử bạn bác bỏ *giả thiết không* là $\beta_2 = 0$. Bạn cũng có thể bác bỏ giả thiết là $\beta_2 = 1$ được không? Bạn sử dụng kiểm định nào cho giả thiết thứ hai?
(h) Bạn có thể kiểm định giả thiết là $\beta_2 = 1$ sử dụng kiểm định F của ANOVA được không? Tại sao hay tại sao không?

5.9. Từ bài tập 3.19

(a) Ước lượng hai hồi quy đưa ra ở đó, tính các kết quả thông thường như các sai số chuẩn, v.v…

(b) Kiểm định giả thiết là các yếu tố nhiễu trong hai mô hình hồi quy có phân phối chuẩn.

(c) Trong hồi quy giá vàng, kiểm định giả thiết là β2 = 1, tức là có mối quan hệ một - một giữa giá vàng và chỉ số giá tiêu dùng - CPI (nghĩa là vàng được bảo đảm hoàn hảo về giá trị). Giá trị p của thống kê kiểm định ước lượng bằng bao nhiêu?

(d) Lập lại bước (c) cho hồi quy chỉ số NYSE (Thị trường Chứng khoán New York). Đầu tư vào thị trường cổ phiếu có được bảo đảm hoàn hảo chống lại lạm phát hay không? Bạn đang kiểm định giả thiết không nào? Giá trị p bằng bao nhiêu?

(e) Giữa vàng và cổ phiếu, bạn lựa chọn loại đầu tư nào? Đâu là cơ sở cho quyết định của bạn?

5.10. Tham khảo Bài tập 3.20. Lập bảng ANOVA để kiểm định giả thiết là thay đổi mức cung tiền không có tác động tới giá tiêu dùng tại Nhật Bản trong khoảng thời gian nghiên cứu.

5.11. Tham khảo Bài tập 3.21.

(a) Giữa sở hữu điện thoại và GDP bình quân đầu người ở Singapore trong thời kỳ 1960-1981 có quan hệ không? Làm sao biết được?
(b) Giả sử GDP thực bình quân đầu người trong năm 1982 là 5.752 USD. Giá trị trung bình ước lượng của Y, số lượng điện thoại trong 1.000 dân, bằng bao nhiêu trong năm đó? Thiết lập một khoảng tin cậy 95% cho ước lượng này.

5.12. Tham khảo Bài tập 1.1. Đối với mỗi quốc gia, ước lượng mô hình sau:

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

với Y_t = tỷ số lạm phát trong thời gian t
X_t = thời gian, lấy giá trị 1, 2,…, 21
u_t = số hạng nhiễu ngẫu nhiên

(a) Bạn có thể đưa ra các kết luận tổng quát nào về tác động của lạm phát tại từng quốc gia?

Bancroft, *Statistical Theory in Research* (Lý thuyết thống kê trong nghiên cứu), McGraw-Hill, New York, 1952, trang 87-88.

(b) Đối với hồi quy của từng quốc gia, kiểm định giả thiết cho rằng β_2, hệ số xu hướng, lớn hơn 0. (Sử dụng mức ý nghĩa 5%)

5.13. Tiếp tục với số liệu trong Bài tập 1.1 và ước lượng hồi quy sau:

$$Y_{it} = \beta_1 + \beta_2 X_t + u_t$$

với Y_{it} = tỷ số lạm phát tại quốc gia i, i là Anh Quốc, Nhật Bản, Đức, hay Pháp.
X_t = tỷ số lạm phát tại Hoa Kỳ.

(a) Đối với từng hồi quy, giữa tỷ số lạm phát của từng quốc gia này với tỷ số lạm phát của Hoa Kỳ có mối quan hệ không?
(b) Bạn tiến hành kiểm định mối quan hệ đó một cách chính thức như thế nào?
(c) Bạn có thể sử dụng để dự báo tỷ số lạm phát của 4 quốc gia sau năm 1980 được không? Tại sao hay tại sao không?

5.14. Bảng sau cung cấp số liệu GNP và bốn định nghĩa về lượng cung tiền đối với Hoa Kỳ trong thời kỳ 1970-1983.

GNP và bốn đại lượng đo lượng cung tiền

	GNP,	Thước đo lượng cung tiền, tỷ USD			
Năm	**tỷ USD**	**M_1**	**M_2**	**M_3**	**L**
1970	992,7	216,6	628,2	677,5	816,3
1971	1.077,6	230,8	712,8	776,2	903,1
1972	1.185,9	252,0	805,2	886,0	1.023,0
1973	1.326,4	265,9	861,0	985,0	1.141,7
1974	1.434,2	277,6	908,5	1.070,5	1.247,3
1975	1.549,2	291,2	1.023,3	1.174,2	1.367,9
1976	1.718,0	310,4	1.163,6	1.311,9	1.516,6
1977	1.918,3	335,4	1.286,7	1.472,9	1.704,7
1978	2.163,9	363,1	1.389,1	1.647,1	1.910,6
1979	2.417,8	389,1	1.498,5	1.804,8	2.117,1
1980	2.631,7	414,9	1.632,6	1.990,0	2.326,2
1981	2.957,8	441,9	1.796,6	2.238,2	2.599,8
1982	3.069,3	480,5	1.965,4	2.462,5	2.870,8
1983	3.304,8	525,4	2.196,3	2.710,4	3.183,1

Định nghĩa:
M_1 = tiền mặt + tiền gửi không kỳ hạn + séc du lịch và các loại tiền gửi được rút séc khác (OCDs)
M_2 = M_1 + Hợp đồng mua lại chứng khoán (RP) 1 ngày đêm và Eurodollar + số dư của MMMF (quỹ hỗ tương trên thị trường tiền tệ) + MMDAs (các tài khoản tiền gửi trên thị trường tiền tệ) + tiết kiệm và tiền gửi nhỏ
M_3 = M_2 + tiền gửi có kỳ hạn lớn + Hợp đồng mua lại chứng khoán có kỳ hạn + MMMF định chế
L = M_3 + các tài sản thanh khoản khác

Nguồn: *Economic Report of the PresidentI* (Báo cáo Kinh tế của Tổng thống), 1985, số liệu GNP lấy từ Bảng B-1, trang 232; số liệu lượng cung tiền lấy từ Bảng B-61, trang 303.

Thực hiện hồi quy GNP theo các định nghĩa về lượng cung tiền khác nhau, ta có các kết quả trong bảng sau:

Hồi quy GNP - lượng cung tiền, 1970-1983

1) GNP_t	=	–787,4723	+	$8{,}6063M_{1t}$	$r^2 = 0{,}9912$
		(77,9664)		(0,2197)	
2) GNP_t	=	–44,0626	+	$1{,}5875\ M_{2t}$	$r^2 = 0{,}9905$
		(61,0134)		(0,0448)	
3) GNP_t	=	159,1366	+	$1{,}2034\ M_{3t}$	$r^2 = 0{,}9943$
		(42,9882)		(0,0262)	
4) GNP_t	=	164,2071	+	$1{,}0290\ L_t$	$r^2 = 0{,}9938$
		(44,7658)		(0,0234)	

Ghi chú: Các số liệu trong ngoặc là sai số chuẩn ước lượng.

Các nhà kinh tế học tiền tệ hay các nhà ủng hộ lý thuyết định lượng cho rằng thu nhập danh nghĩa (nghĩa là GNP danh nghĩa) được xác định phần nhiều bởi thay đổi lượng cung tiền, mặc dù không có sự nhất trí trong định nghĩa "đúng" về tiền tệ. Với các kết quả ở bảng trên, trả lời các câu hỏi sau:

(a) Định nghĩa nào về tiền tệ có vẻ có quan hệ mật thiết với GNP danh nghĩa?
(b) Do đại lượng r^2 có giá trị cao một cách đồng nhất, có phải nó có nghĩa là sự lựa chọn định nghĩa tiền tệ của chúng ta không làm thay đổi gì cả?
(c) Nếu Hệ thống Dự trữ Liên bang (Fed) muốn kiểm soát lượng cung tiền, đại lượng đo tiền tệ nào là đối tượng tốt nhất cho mục tiêu này? Bạn có thể suy ra từ các kết quả hồi quy được không?

5.15. Giả sử phương trình của một **đường đẳng dụng** giữa hai hàng hóa là

$$X_iY_i = \beta_1 + \beta_2 X_i$$

Bạn làm thế nào để ước lượng các thông số của mô hình này? Áp dụng mô hình trên cho các số liệu sau và bình luận các kết quả của bạn:

Tiêu dùng hàng hóa *X*	1	2	3	4	5
Tiêu dùng hàng hóa *Y*	4	3,5	2,8	1,9	0,8

5.16. Đường thị trường vôn (CML) của lý thuyết đầu tư chứng khoán* mặc định một mối quan hệ tuyến tính giữa suất sinh lợi kỳ vọng và rủi ro (tính bằng độ lệch chuẩn) cho cơ cấu đầu tư chứng khoản hiệu quả như sau:

$$E_i = \beta_1 + \beta_2\sigma_i$$

với E_i = suất sinh lợi kỳ vọng của chứng khoán i và σ_i = độ lệch chuẩn của suất sinh lợi. Sau đây là số liệu về suất sinh lợi kỳ vọng và độ lệch chuẩn của suất sinh lợi của cơ cấu chứng khoán của 34 quỹ hỗ tương (quỹ đầu tư chung) tại Hoa Kỳ trong giai đoạn 1954-1963. Kiểm tra xem số liệu có hỗ trợ lý thuyết hay không.

5.17. Tham khảo bài tập 3.22. Sử dụng số liệu trong đó, ước lượng mô hình đề xuất đối với GDP tính theo đô la hiện hành trong giai đoạn 1972-1986. Sử dụng mô hình ước lượng, tính các giá trị dự báo của GDP tính theo giá hiện hành cho các năm 1987, 1988, 1989, 1990, 1991 và so sánh chúng với các giá trị thực tế.

* Xem William F. Sharpe, *Portfolio Theory and Capital Markets* (Lý thuyết đầu tư chứng khoán và thị trường vốn), McGraw-Hill, New York, 1970, trang 83.

Kết quả hoạt động của 34 quỹ hỗ tương, 1954-1963

Tên quỹ	Suất sinh lợi bình quân năm (%)	Độ lệch chuẩn của suất sinh lợi hàng năm (%)
Affiliated Fund	14,6	15,3
American Business Shares	10,0	9,2
Axe-Houghton, Fund A	10,5	13,5
Axe-Houghton, Fund B	12,0	16,3
Axe-Houghton, Stock Fund	11,9	15,6
Bosten Fund	12,4	12,1
Board Street Investing	14,8	16,8
Bullock Fund	15,7	19,3
Commonwealth Investment Company	10,9	13,7
Delaware Fund	14,4	21,4
Dividend Shares	14,4	15,9
Eaton and Howard Balanced Fund	11,0	11,9
Eaton and Howard Stock Fund	15,2	19,2
Equity Fund	14,6	18,7
Fidelity Fund	16,4	23,5
Financial Industrial Fund	14,5	23,0
Fundamental Investors	16,0	21,7
Group Securities. Common Stock Fund	15,1	19,1
Group Securities. Fully Administered Fund	11,4	14,1
Incorporated Investors	14,0	25,5
Investment Company of America	17,4	21,8
Investors Mutual	11,3	12,5
Loomis-Sales Mutual Fund	10,0	10,4
Massachusetts Investors Trust	16,2	20,8
Massachusetts Investors-Growth Stock	18,6	22,7
National Investors Corporation	18,3	19,9
National Securities-Income Series	12,4	17,8
New England Fund	10,4	10,2
Putnam Fund of Boston	13,1	16,0
Scudder, Stevens & Clark Balanced Fund	10,7	13,3
Selected American Shares	14,4	19,4
United Funds-Income Fund	16,1	20,9
Wellington Fund	11,3	12,0
Wisconsin Fund	13,8	16,9

Nguồn: William F. Sharpe, "Mutual Fund Performance", *Journal of Business* (Kết quả hoạt động của các quỹ hỗ tương, Tạp chí Kinh doanh), tháng 1 năm 1966, phụ trương, trang 125.

5.18. Từ năm 1986, Tạp chí the *Economist* (Nhà kinh tế) xuất bản Chỉ số Big Mac như là một đại lượng đo thô và vui nhộn để xem các đồng tiền quốc tế có ở vào tỷ giá hối đoái "đúng" của chúng không, như theo lý thuyết **cân bằng sức mua (PPP)**. Lý thuyết PPP cho rằng một đơn vị tiền tệ phải có khả năng mua cùng một nhóm hàng hóa ở tất cả các nước. Những người đề xuất PPP lập luận rằng, trong thời gian dài hạn, các đồng tiền có xu hướng dịch chuyển về giá trị PPP của chúng. Tạp chí the *Economist* sử dụng chỉ số Big Mac của McDonald như là một nhóm đại diện và đưa ra thông tin sau.

Tiêu chuẩn bánh mì kẹp (hamburger)

	Giá Big Mac		Tỷ giá hối đoái thực tế 5/4/94	PPP† quy cho của USD	Phá giá (-)/Nâng giá (+)** Nội tệ
	bằng nội tệ*	bằng USD			
UNITED STATES	USD2,30	2,30	–	–	–
Argentina	Peso3,60	3,60	1,00	1,57	+57
Úc	A$2,45	1,72	1,42	1,07	-25
Áo	Scl34,00	2,84	12,00	14,80	+23
Bỉ	BFr109	3,10	35,20	47,39	-35
Brazil	Crl,500	1,58	949,00	652,00	-31
Anh Quốc	£1,81	2,65	1.46‡‡	1.27‡‡	+15
Canada	C$2,86	2,06	1,39	1,24	-10
Chilê	Peso948	2,28	414,00	412,00	-1
Trung Quốc	Yuan9,00	1,03	8,70	3,91	-55
CH Séc	CKr50	1,71	29,70	217,00	-27
Đan Mạch	DKr25,75	3,85	6,69	11,20	+67
Pháp	FFr 18,5	3,17	5,83	8,04	+38
Đức	DM4,60	2,69	1,71	2,00	+17
Thụy Sĩ	Dr620	2,47	251,00	270,00	+8
Hà Lan	F15,45	2,85	1,91	2,37	+24
Hồng Kông	HK$9,20	1,19	7,73	4,00	-48
Hungary	Forintl69	1,66	103,00	73,48	-29
Ý	Lire4,550	2,77	1.641,00	1.978,00	+21
Nhật Bản	V391	3,77	104,00	170,00	+64
Malaysia	M$3,77	1,40	2,69	1,64	-39
Mêhicô	Peso8,1 0	2,41	3,36	3,52	+5
Ba Lan	Zloty31.000	1,40	2,69	13.478,00	-40
Bồ Đào Nha	Esc440	2,53	174,00	191,00	+10
Nga	Rouble2.900	1,66	17.775,00	1.261,00	-29
Singapore	$2,98	1,90	1,57	1,30	-17
Hàn Quốc	Won2.300	2,84	810,00	1.000,00	+24
Tây Ban Nha	Ptas345	2,50	138,00	150,00	+9
Thụy Điển	Skr25,5	3,20	7,97	11,10	+39
Thụy Sĩ	SFr5,70	3,96	1,44	2,48	+72
Đài Loan	NT$62	2,35	26,40	26,96	+2
Thái Lan	Baht48	1,90	25,30	20,87	-17

* Giá cả thay đổi trong nội địa.
† Cân bằng sức mua: giá nội địa chia cho giá tại Hoa Kỳ.
** So với đô la.
‡ Trung bình của New York, Chicago, San Francisco, và Atlanta.
‡‡ Đô la/pound.

Nguồn: McDonald và Tạp chí *The Economist*, 9/4/1994, trang 88.

Xem xét mô hình hồi quy sau:

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

với Y = tỷ giá hối đoái thực tế và X = PPP quy cho của đô la.

(a) Nếu PPP đúng, bạn có một tiên nghiệm gì về giá trị của β_1 và β_2?
(b) Các kết quả hồi quy có hỗ trợ ước đoán của bạn không? Bạn sử dụng kiểm định chính thức nào để kiểm định giả thiết của mình?
(c) Tạp chí the *Economist* có nên tiếp tục xuất bản Chỉ số Big Mac không? Tại sao hay tại sao không?

5.19. Tham khảo số liệu S.A.T trong bài tập 2.16. Giả sử bạn muốn dự đoán điểm toán của học sinh nam (Y) trên cơ sở điểm toán của học sinh nữ (X) bằng cách chạy hồi quy sau:

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

(a) Ước lượng mô hình trên.
(b) Từ các phần dư ước lượng, xác định xem giả thiết về quy luật chuẩn có đứng vững không?
(c) Bây giờ, kiểm định giả thiết là $\beta_2 = 1$, tức là, có sự tương ứng một - một giữa điểm toán của học sinh nam và nữ.
(d) Thiết lập bảng ANOVA cho bài tập này.

5.20. Lặp lại bài tập trên nhưng với Y và X đại diện cho điểm vấn đáp tương ứng của học sinh nam và nữ.

PHỤ LỤC 5A

5A.1 NGUỒN GỐC PHƯƠNG TRÌNH (5.3.2)

Đặt

$$Z_1 = \frac{\hat{\beta}_2 - \beta_2}{se(\hat{\beta}_2)} = \frac{(\hat{\beta}_2 - \beta_2)\sqrt{\sum x_i^2}}{\sigma} \qquad (1)$$

và

$$Z_2 = (n-2)\frac{\hat{\sigma}^2}{\sigma^2} \qquad (2)$$

Với σ đã biết, Z_1 tuân theo phân phối chuẩn hóa; tức là $Z_1 \sim N(0, 1)$. (Tại sao?) Z_2 tuân theo phân phối χ^2 với (n – 2) bậc tự do. (Để chứng minh xem chú thích 5). Hơn nữa, ta có thể chỉ ra rằng Z_2 có phân phối độc lập với Z_1.* Do vậy, theo Định lý 4.5, biến số

$$t = \frac{Z_1\sqrt{n-2}}{\sqrt{Z_2}} \qquad (3)$$

tuân theo phân phối t với n – 2 bậc tự do. Thay thế (1) và (2) vào (3), ta có phương trình 5.3.2).

5A.2 NGUỒN GỐC PHƯƠNG TRÌNH (5.9.1)

Phương trình (1) cho biết $Z_1 \sim N(0, 1)$. Do vậy, theo Định lý 4.3,

$$Z_1^2 = \frac{(\hat{\beta}_2 - \beta_2)^2 \sum x_i^2}{\sigma^2}$$

* Để chứng minh, Xem J. Johnston, Econometric Methods (Các phương pháp Kinh tế lượng), MacGraw-Hill, in lần thứ 3, NewYork, 1984, trang 181-182 (Cần có kiến thức đại số ma trận để theo dõi chứng minh này)

tuân theo phân phối χ^2 với 1 bậc tự do. Như đã đề cập trong Mục 5A.1,

$$Z_2 = (n-2)\frac{\hat{\sigma}^2}{\sigma^2} = \frac{\sum \hat{u}_i^2}{\sigma^2}$$

cũng tuân theo phân phối χ^2 với n – 2 bậc tự do. Hơn nữa, như đã lưu ý trong Mục 4.3., Z_2 có phân phối độc lập với Z_1. Do vậy, áp dụng Định lý 4.6, ta có

$$F = \frac{Z_1^2/1}{Z_2/(n-2)} = \frac{(\hat{\beta}_2 - \beta_2)^2(\sum x_i^2)}{\sum \hat{u}_i^2/(n-2)}$$

tuân theo phân phối F tương ứng với 1 và n – 2 bậc tự do. Theo giả thiết không H_0: $\beta_2 = 0$, tỷ lệ F ở trên rút gọn thành phương trình (5.9.1).

5.A.3. NGUỒN GỐC PHƯƠNG TRÌNH (5.10.2) VÀ (5.10.6)

Phương sai của dự báo giá trị trung bình

Với $X_i = X_0$, giá trị đúng của dự báo trung bình $E(Y_0 \mid X_0)$ được tính bởi:

$$E(Y_0 \mid X_0) = \beta_1 + \beta_2 X_0 \tag{1}$$

Ta ước lượng (1) từ

$$\hat{Y}_0 = \beta_1 + \beta_2 X_0 \tag{2}$$

Lấy kỳ vọng của (2), với X_0 cho trước, ta có

$$E(\hat{Y}_0) = E(\hat{\beta}_1) + E(\hat{\beta}_2)X_0$$
$$= \beta_1 + \beta_2 X_0$$

Bởi vì $\hat{\beta}_1$ và $\hat{\beta}_2$ là các ước lượng không thiên lệch, Vì vậy

$$E(\hat{Y}_0) = E(Y_0 \mid X_0) = \beta_1 + \beta_2 X_0 \tag{3}$$

Tức là, $\hat{Y}_0$ là một đại lượng dự báo không thiên lệch của $E(Y_0 \mid X_0)$.

Bây giờ bằng cách sử dụng tính chất

Var (a+b) = Var (a) + Var (b) +2 Covar (a,b)

Chúng ta thu được

$$\text{Var}(\hat{Y}_0) = \text{Var}(\hat{\beta}_1) + \text{Var}(\hat{\beta}_2)X_0^2 + 2\,\text{Covar}(\hat{\beta}_1, \hat{\beta}_2)X_0$$

Bây giờ, sử dụng các công thức về phương sai và đồng phương sai (tích sai) của $\hat{\beta}_1$ và $\hat{\beta}_2$ trong (3.3.1), (3.3.3) và (3.3.9) và biến đổi các số hạng, ta có

$$\text{var}(\hat{Y}_0) = \sigma^2\left[\frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum x_i^2}\right] = (5.10.2)$$

Phương sai của dự báo điểm cá biệt

Ta muốn dự báo giá trị cá biệt Y tương ứng với $X = X_0$, tức là, ta muốn tính được

$$Y_0 = \beta_1 + \beta_2 X_0 + u_0 \tag{5}$$

Ta dự báo nó như sau

$$\hat{Y}_0 = \hat{\beta}_1 + \hat{\beta}_2 X_0 \tag{6}$$

Sai số dự báo, $Y_0 - \hat{Y}_0$, là

$$\begin{aligned} Y_0 - \hat{Y}_0 &= \beta_1 + \beta_2 X_0 + u_0 - (\hat{\beta}_1 + \hat{\beta}_2 X_0) \\ &= (\beta_1 - \hat{\beta}_1) + (\beta_2 - \hat{\beta}_2)X_0 + u_0 \end{aligned} \tag{7}$$

Do vậy,

$$\begin{aligned} E(Y_0 - \hat{Y}_0) &= E(\beta_1 - \hat{\beta}_1) + E(\beta_2 - \hat{\beta}_2)X_0 - E(u_0) \\ &= 0 \end{aligned}$$

do $\hat{\beta}_1$, $\hat{\beta}_2$ không thiên lệch, X_0 là số cố định, và $E(u_0)$ bằng 0 theo giả thiết.

Bình phương hai vế của (7) và lấy kỳ vọng, ta có $\mathrm{var}(Y_0 - \hat{Y}_0) = \mathrm{var}(\hat{\beta}_1) + X_0^2\,\mathrm{var}(\hat{\beta}_2) + 2X_0\,\mathrm{Cov}(\beta_1, \beta_2) + \mathrm{var}(u_0)$. Sử dụng các công thức phương sai và tích sai cho $\hat{\beta}_1$ và $\hat{\beta}_2$ thiết lập trước đây, và lưu ý rằng $\mathrm{var}(u_0) = \sigma^2$, ta có

$$\mathrm{var}(Y_0 - \hat{Y}_0) = \sigma^2\left[1 + \frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum x_i^2}\right] \qquad = (5.10.6)$$