

**BỘ CÔNG THƯƠNG**  
**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP QUẢNG NINH**



**GIÁO TRÌNH**  
**KINH TẾ LƯỢNG**

**DÙNG CHO BẬC ĐẠI HỌC**  
**(LƯU HÀNH NỘI BỘ)**

**QUẢNG NINH - 2020**

## KINH TẾ LƯỢNG

**1. Tên học phần:** Kinh Tế Lượng (Econometrics)

**2. Mô tả vắn tắt nội dung học phần:**

Mục đích của Kinh Tế Lượng là giúp học viên nắm rõ và vận dụng được các mô hình hồi qui để ước lượng, dự đoán giá trị trung bình của tổng thể của các biến phụ thuộc theo giá trị của biến giải thích nhằm xác định mức độ quan hệ giữa các biến, từ đó thấy được bản chất của các hiện tượng và tìm được các biện pháp khắc phục. Môn học còn nhằm trang bị cho các học viên cách thức vận dụng các công cụ phân tích định lượng vào việc xử lý phân tích các vấn đề kinh tế cụ thể.

**3. Nhiệm vụ của học viên:**

Dự giờ giảng trên lớp và đọc giáo trình, làm bài tập theo nhóm về xử lý dữ liệu bằng ít nhất một phần mềm được môn học trang bị, tham dự thảo luận dưới sự hướng dẫn của giảng viên. Tham dự kiểm tra hết môn học theo lịch nhà trường qui định.

**4. Tài liệu tham khảo thêm:**

- Basic Econometrics, tác giả Damodar N. Gujarati, 1995.
- Introductory Econometrics, tác giả Craig A. Depken, 2006
- Econometric Analysis, tác giả William H . Greene, 2000.

**5. Tiêu chuẩn đánh giá:**

- Dự giờ đủ trên lớp theo yêu cầu của môn học
- Hoàn thành và đạt yêu cầu các bài tập về môn học trước khi thi hết môn
- Thi hết môn

**6. Mục tiêu của học phần:**

Nắm vững các mô hình kinh tế lượng để có thể lượng hoá các quan hệ kinh tế vĩ mô và vi mô được trang bị trước đây. Liên kết được các mô hình kinh tế lượng với các lý thuyết kinh tế vĩ mô và vi mô bằng các dữ liệu thực tế.

Đề xuất chính sách và dự báo dựa trên việc phân tích, kiểm định các mối quan hệ kinh tế vi mô và vĩ mô qua kết quả của mô hình khi ứng dụng số liệu thực tế.

**7. Nội dung học phần:**

- Chương I: Cơ bản về Kinh tế lượng và phân tích hồi qui
- Chương II: Mô hình hồi qui hai biến, ước lượng và kiểm định
- Chương III: Mô hình hồi qui đa biến
- Chương IV: Đa cộng tuyến
- Chương V: Hồi qui với biến giả
- Chương VI: Phương sai sai số thay đổi
- Chương VII: Tương quan chuỗi

## CHƯƠNG I

### CƠ BẢN VỀ KINH TẾ LƯỢNG VÀ PHÂN TÍCH HỒI QUI

#### 1.1. Vài nét cơ bản về kinh tế lượng:

##### 1.1.1. Giới thiệu về kinh tế lượng

Thuật ngữ tiếng Anh là Econometrics, nó được ghép từ 2 từ “Econo” có nghĩa là kinh tế và “Metrics” có nghĩa là đo lường- Vậy đó là “đo lường kinh tế”.

Theo nghĩa đơn giản, **kinh tế lượng**, liên quan đến việc áp dụng các phương pháp thống kê trong kinh tế. Không như *thống kê kinh tế*, trong đó các dữ liệu thống kê là chính yếu, kinh tế lượng được phân biệt bằng sự *hợp nhất* của lý thuyết kinh tế, công cụ toán học và các phương pháp luận thống kê. Mở rộng hơn, kinh tế lượng quan tâm đến (1) ước lượng các mối quan hệ kinh tế, (2) đối chiếu lý thuyết kinh tế với thực tế và kiểm định các giả thuyết liên quan đến hành vi kinh tế, và (3) dự báo các hành vi của các biến số kinh tế.

Người ta có thể định nghĩa như sau:

- + Kinh tế lượng bao gồm việc áp dụng thống kê toán cho các số liệu kinh tế để củng cố về mặt thực nghiệm cho các mô hình do các nhà kinh tế toán đề xuất và để tìm ra lời giải bằng số.
- + Kinh tế lượng có thể được định nghĩa như là sự phân tích về lượng các vấn đề kinh tế hiện thời, dựa trên việc vận dụng đồng thời lý thuyết và thực tế được thực hiện bằng các phương pháp suy đoán thích hợp.

Ví dụ về ứng dụng của kinh tế lượng trong:

#### Ước lượng các mối quan hệ kinh tế

Kinh tế học thực nghiệm cung cấp rất nhiều ví dụ nhằm ước lượng các mối quan hệ kinh tế như:

1. Ước lượng cầu/cung của các sản phẩm, dịch vụ.
2. Ước lượng ảnh hưởng của chi phí bán hàng/quảng cáo đến doanh thu và lợi nhuận.
3. Giá của cổ phiếu với các đặc trưng của công ty phát hành cổ phiếu đó, cũng như với tình hình chung của nền kinh tế.
4. Đánh giá tác động của các chính sách tiền tệ và tài chính đến các biến như việc làm hoặc thất nghiệp, thu nhập, xuất khẩu và nhập khẩu, lãi suất, tỷ lệ lạm phát, và thâm hụt ngân sách.

#### Kiểm định giả thuyết

Cũng như bất kỳ ngành khoa học nào, một ưu điểm của kinh tế lượng là quan tâm đến **việc kiểm định giả thuyết** về các hành vi kinh tế. Ví dụ như:

1. Một doanh nghiệp có thể muốn xác định xem chiến dịch quảng cáo của mình có tác động làm tăng doanh thu hay không.
2. Các nhà phân tích có thể quan tâm xem nhu cầu có giãn hay không co giãn theo giá và thu nhập.
3. Công ty muốn biết lợi nhuận có tăng hay giảm theo qui mô hoạt động không.
4. Các nhà kinh tế học vĩ mô có thể muốn đánh giá hiệu quả của các chính sách nhà nước.

### **Dự báo**

Khi các biến số được xác định và chúng ta đánh giá được tác động cụ thể của chúng đến chủ thể nghiên cứu, chúng ta có thể muốn sử dụng các mối quan hệ ước lượng để dự đoán các giá trị trong tương lai.

Ví dụ:

1. Các công ty dự báo doanh thu, lợi nhuận, chi phí sản xuất, và lượng tồn kho cần thiết.
2. Dự đoán có nhu cầu về năng lượng nhằm phục vụ việc hoạch định các chính sách có liên quan.
3. Dự báo các chỉ số thị trường chứng khoán và giá của một số cổ phiếu.
4. Dự đoán thu nhập, chi tiêu, lạm phát, thất nghiệp, và thâm hụt ngân sách và thương mại.
5. Các thành phố dự báo định kỳ mức tăng trưởng của địa phương qua các mặt như: dân số; việc làm; số nhà ở, nhu cầu về trường học, và dịch vụ công cộng; ...v.v

#### **1.1.2. Mục đích của kinh tế lượng**

Mục đích của kinh tế lượng là giải thích sự biến thiên của biến và các mối quan hệ của các biến, ví dụ:

Có 1 biến (chỉ tiêu) thay đổi (do lệch khỏi trung bình) mà chúng ta cần phải giải thích, ví dụ khi chúng ta nghiên cứu lượng bán một loại sản phẩm nào đó (Q) biến động, vậy cái gì tác động đến nó và các chỉ tiêu tác động lẫn nhau như thế nào.

#### **1.1.3. Phương pháp luận của kinh tế lượng**

Nêu ra các giả thuyết hay giả thiết về các mối quan hệ giữa các biến kinh tế: chẳng hạn kinh tế vĩ mô khẳng định rằng mức tiêu dùng của các hộ gia đình phụ thuộc theo quan hệ cùng chiều với thu nhập khả dụng của họ.

Thiết lập mô hình toán học để mô tả mối quan hệ giữa các biến số này Các phương trình này mô tả mối quan hệ giữa các biến số kinh tế với nhau. Một phương trình sẽ bao gồm một biến phụ thuộc và một hoặc nhiều biến giải thích. Sự tác động của một biến giải thích lên biến phụ thuộc được đo lường bằng hệ số của nó và hình thức hàm của phương trình. Một phương trình tiêu biểu như sau:

$$Y(t) = f\{x_1(t), x_2(t), \dots, x_n(t), u(t)\}$$

$Y(t)$  là biến phụ thuộc tại thời điểm  $t$ , biểu trưng cho chỉ tiêu cần nghiên cứu hay dự báo (ví dụ như GDP, việc làm, lạm phát,...).

$x_1(t), x_2(t), \dots, x_n(t)$  là các biến giải thích tại thời điểm  $t$ , biểu trưng cho các nhân tố tác động lên biến phụ thuộc. Sự thay đổi của một hay nhiều biến này sẽ dẫn tới sự thay đổi của biến phụ thuộc.

$u(t)$  là sai số ngẫu nhiên, biểu trưng cho các nhân tố không xác định được tác động lên biến phụ thuộc tại thời điểm  $t$ .

Số hạng sai số  $u(t)$ , chúng ta cũng có thể ký hiệu là  $u_i$  (hay còn gọi là số hạng nhiễu ngẫu nhiên) là thành phần ngẫu nhiên không quan sát được và là sai biệt giữa  $Y_i$  và phần xác định  $\beta_1 + \beta_2 X_i$ .

Sau đây một tổ hợp của bốn nguyên nhân ảnh hưởng khác nhau:

1. *Biến bỏ sót.* Giả sử mô hình thực sự là  $Y_i = \beta_1 + \beta_2 X_i + \beta_3 Z_i + v_i$  trong đó,  $Z_i$  là một biến giải thích khác và  $v_i$  là số hạng sai số thực sự, nhưng nếu ta sử dụng mô hình là  $Y_i = \beta_1 + \beta_2 X_i + u_i$  thì  $u_i = \beta_3 Z_i + v_i$ . Vì thế,  $u_i$  bao hàm cả ảnh hưởng của biến  $Z$  bị bỏ sót.
2. *Phi tuyến tính.*  $u_i$  có thể bao gồm ảnh hưởng phi tuyến tính trong mối quan hệ giữa  $Y$  và  $X$ . Vì thế, nếu mô hình thực sự là  $Y_i = \beta_1 + \beta_2 X_i + \beta_3 X_i^2 + v_i$ , nhưng lại được giả định bằng phương trình  $Y_i = \beta_1 + \beta_2 X_i + u_i$ , thì ảnh hưởng của  $X_i^2$  sẽ được bao hàm trong  $u_i$ .
3. *Sai số đo lường.* Sai số trong việc đo lường  $X$  và  $Y$  có thể được thể hiện thông qua  $u$ .
4. *Những ảnh hưởng không thể dự báo.* Dù là một mô hình kinh tế lượng tốt cũng có thể chịu những ảnh hưởng ngẫu nhiên không thể dự báo được. Những ảnh hưởng này sẽ luôn được thể hiện qua số hạng sai số  $u_i$ .

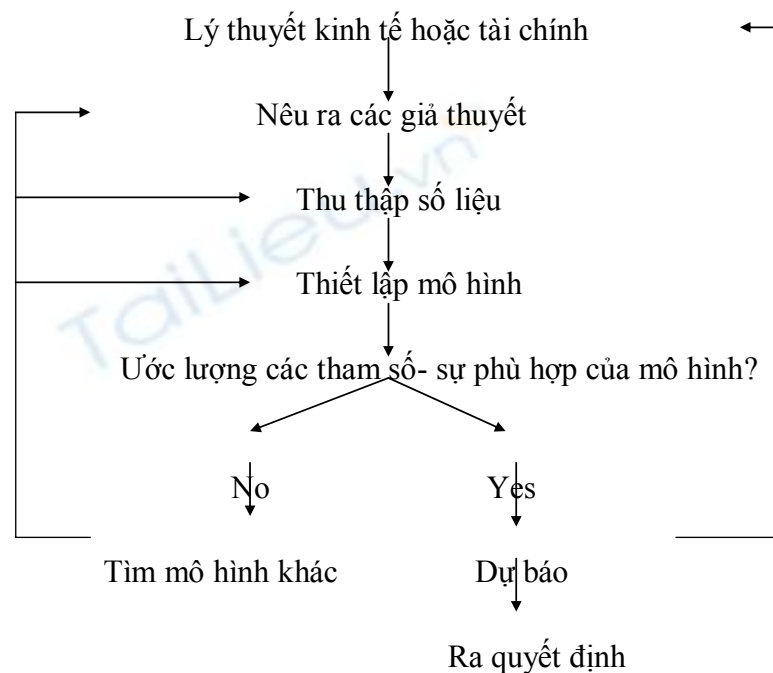
Việc xây dựng hệ thống các phương trình, với các biến giải thích lựa chọn thường được dựa trên nền tảng của lý thuyết kinh tế. Ví dụ như hàm tiêu dùng phải dựa trên lý thuyết về tiêu dùng, hàm đầu tư phải dựa trên lý thuyết về đầu tư,... Điều này dẫn đến hệ quả là các nhà mô hình khác nhau có thể sẽ xây dựng các phương trình với các biến giải thích khác nhau, tùy thuộc vào việc áp dụng lý thuyết kinh tế nào. Điều đó cũng lý giải về sự đa dạng của các mô hình kinh tế lượng hiện nay.

Ví dụ, Giả sử chúng ta điều tra tất cả các hộ trong thành phố và tính thu nhập hàng tháng của họ ( $X$ ) và tổng chi tiêu vào hàng hóa và dịch vụ ( $Y$ ). Vì các hộ gia đình có cùng thu nhập sẽ có những mức chi tiêu khác nhau (có lẽ do khác biệt về các đặc điểm khác như số thành viên trong gia đình), một quan sát cụ thể ( $Y, X$ ) sẽ không hoàn toàn chính xác nằm trên đường thẳng. Do vậy, mô hình hồi qui tuyến tính tương ứng với ví dụ này sẽ có dạng  $Y = \beta_1 + \beta_2 X + u$

Trong thực tế, chúng ta sẽ không điều tra tất cả các hộ gia đình mà chỉ chọn một mẫu ngẫu nhiên từ tổng thể và sử dụng các quan sát này để ước lượng những tham số  $\beta_1$  và  $\beta_2$  cũng như thực hiện các kiểm định và kiểm tra tính phù hợp của giả định về mối liên hệ trung bình giữa chi tiêu và thu nhập là tuyến tính.

Sau khi xây dựng xong hệ thống các phương trình, chúng ta phải tập hợp đủ các số liệu cho các biến và tiến hành ước lượng các hệ số của các phương trình. Kỹ thuật hồi quy (regression) được áp dụng để ước lượng các hệ số của các phương trình. Sau khi ước lượng xong toàn bộ các phương trình của mô hình, chúng ta sẽ tiến hành mô phỏng (simulation) tác động của các thay đổi chính sách trong tương lai lên các biến kinh tế mà mình quan tâm (ví dụ như tăng trưởng, việc làm, lạm phát,...). Trên cơ sở đó, chúng ta có thể đánh giá tác động của chúng hoặc/và đề xuất ra các kịch bản dự báo.

### Các bước thực hiện



## 1.2. Phân tích hồi qui

### 1.2.1. Các ví dụ trong lĩnh vực kinh tế về mối quan hệ nhân quả

Trong phân tích hồi qui, chúng ta cần ước lượng quan hệ toán học giữa các biến. Những mối quan hệ này còn được gọi là mối quan hệ hàm số. Chúng cố gắng mô tả các biến giải thích tác động lên biến phụ thuộc như thế nào.

- Biến giải thích là biến xảy ra.
- Biến phụ thuộc là biến kết quả.

Ví dụ: Khi chúng ta cố gắng giải thích chi tiêu dùng của mọi người, chúng ta có thể sử dụng các biến giải thích là thu nhập và độ tuổi. Khi giải thích giá của một chiếc ô tô, các biến giải thích có thể là kích cỡ, động cơ máy, độ tin cậy của hãng sản xuất cũng như độ an toàn của chiếc ô tô. Để giải thích giá của một ngôi nhà các biến giải thích có thể là kích cỡ, số phòng, tỷ lệ tội phạm của khu dân cư cũng như độ tuổi của ngôi nhà. Để dự đoán khả năng một học sinh cuối cấp trung học phổ thông vào đại học, chúng ta có thể

xem xét đến điểm các bài kiểm tra, trình độ giáo dục của cha mẹ cũng như thu nhập của gia đình anh ta.

Vậy phân tích hồi qui chính là nghiên cứu mối liên hệ phụ thuộc của một biến (gọi là biến phụ thuộc hay biến được giải thích) với một hay nhiều biến khác (được gọi là biến độc lập hay giải thích).

### **1.2.2. Mục đích của phân tích hồi qui:**

Tưởng tượng rằng chúng ta có thông tin về thu nhập và chi tiêu tiêu dùng, chúng ta tin tưởng rằng chi tiêu tiêu dùng phụ thuộc vào thu nhập và chúng ta biểu diễn cả 2 biến này lên đồ thị. Biểu diễn biến phụ thuộc lên trục tung, còn biến giải thích (biến độc lập) lên trục hoành.

Mục đích của phân tích hồi qui là qua những điểm dữ liệu, chúng ta có thể kẻ ra một đường phù hợp nhất, sát nhất với các quan sát để sao cho có thể biểu diễn mối quan hệ giữa hai biến thu nhập và chi tiêu tiêu dùng một cách đáng tin cậy nhất.

### **1.2.3. Giới thiệu mô hình hồi qui tuyến tính đơn giản:**

Để mô hình hóa quan hệ tuyến tính trong đó diễn tả sự thay đổi của biến Y theo biến X cho trước người ta sử dụng mô hình hồi qui tuyến tính đơn giản. Mô hình hồi qui tuyến tính đơn giản có dạng sau:

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

- +  $Y_i$  : Giá trị của biến phụ thuộc Y trong lần quan sát thứ i.
- +  $X_i$  : Giá trị của biến độc lập X trong lần quan sát thứ i .
- +  $u_i$  : Giá trị đối với sự dao động ngẫu nhiên (nhiều ngẫu nhiên) hay sai số trong lần quan sát thứ i.
- +  $\beta_1$  : là thông số diễn tả tung độ gốc (hệ số chặn) của đường hồi qui tổng thể, hay  $\beta_1$  là giá trị trung bình của biến phụ thuộc Y khi biến độc lập X thay đổi 1 đơn vị.
- +  $\beta_2$  : là thông số diễn tả độ dốc (hệ số góc) của đường hồi qui tổng thể, hay  $\beta_2$  diễn tả sự thay đổi của giá trị trung bình của biến phụ thuộc Y khi biến độc lập X thay đổi 1 đơn vị.

Chúng ta có thể ước lượng các tham số ( $\beta_1, \beta_2$ ) của phương trình hồi qui tổng thể bằng cách sử dụng số liệu của mẫu ngẫu nhiên thu thập được. Dựa vào số liệu của mẫu ta có phương trình hồi qui tuyến tính của mẫu.

$$\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i}$$

Trong đó:

$\hat{Y}$  là ước lượng của giá trị trung bình của Y đối với biến X đã biết

$\hat{\beta}_1$  là ước lượng của  $\beta_1$

$\hat{\beta}_2$  là ước lượng của  $\beta_2$

### 1.3. Hồi qui và tương quan

Khi định mô hình ở dạng  $Y_i = \beta_1 + \beta_2 X_i + u_i$ , chúng ta ngầm giả định rằng  $X$  gây ra sự thay đổi của  $Y$ . Việc  $X$  và  $Y$  tương quan chặt với nhau không có nghĩa rằng sự thay đổi  $X$  dẫn đến sự thay đổi  $Y$  hay ngược lại. Ví dụ, hệ số tương quan giữa số lượng kanguru của Úc và tổng dân số nước này có thể là rất cao. Phải chăng điều này có nghĩa rằng sự thay đổi một biến sẽ làm cho biến kia thay đổi? Rõ ràng là không, vì ở đây chúng ta có một trường hợp **tương quan giả tạo**. Nếu chúng ta hồi quy một trong các biến với biến còn lại, chúng ta sẽ có sự **hồi qui giả tạo**. Lấy một ví dụ khác thực tế hơn, giả sử chúng ta hồi quy số lượng vụ trộm trong một thành phố với số hạng hằng số và số nhân viên cảnh sát ( $X$ ) và sau đó quan sát thấy hệ số góc ước lượng có giá trị dương, có nghĩa rằng có tương quan thuận giữa  $X$  và  $Y$ . Phải chăng điều này có nghĩa rằng việc tăng số lượng cảnh sát sẽ làm tăng số vụ trộm, do đó ngầm kéo theo phải có chính sách giảm lực lượng cảnh sát? Rõ ràng kết luận này là không thể chấp nhận được. Điều xảy ra có thể là mối quan hệ nhân quả là ngược lại, có nghĩa là thành phố nên thuê thêm cảnh sát vì số vụ trộm tăng lên, và như vậy việc hồi quy  $X$  theo  $Y$  là hợp lý hơn.

Từ những ví dụ trên ta thấy rằng Hồi qui và tương quan khác nhau về mục đích và kỹ thuật. Phân tích tương quan xem xét mức độ kết hợp tuyến tính giữa hai biến, nhưng phân tích hồi qui lại ước lượng hoặc dự báo một biến trên cơ sở giá trị đã cho của các biến khác.

Về mặt kỹ thuật, trong phân tích hồi qui các biến không có tính chất đối xứng, biến phụ thuộc là biến ngẫu nhiên, các biến giải thích giá trị của chúng đã được xác định. Trong phân tích tương quan, không có sự phân biệt giữa các biến, chúng có tính chất đối xứng.

### 1.4. Các dạng hàm trong kinh tế lượng

Giả sử ta có một mô hình kinh tế tiên đoán mối quan hệ giữa một biến phụ thuộc  $Y$  và một biến độc lập  $X$ . Trong nhiều trường hợp, mô hình này sẽ không cho chúng ta biết dạng hàm mà mối quan hệ này có trong dữ liệu, mặc dù mô hình này sẽ thường cho thấy một số ý niệm về dạng có thể có của mối quan hệ. Giải pháp thông thường là quyết định xem dạng hàm nào có khả năng mô tả tốt dữ liệu nhất, điều này hoặc phụ thuộc vào suy luận kinh tế hoặc phụ thuộc vào việc khảo sát dữ liệu. Sau đó, chúng ta thử xây dựng một số dạng hàm khác nhau và xem chúng có cho ra các kết quả tương tự hay không, và nếu không, thì phải xem dạng hàm nào cho ra các kết quả hợp lý nhất.

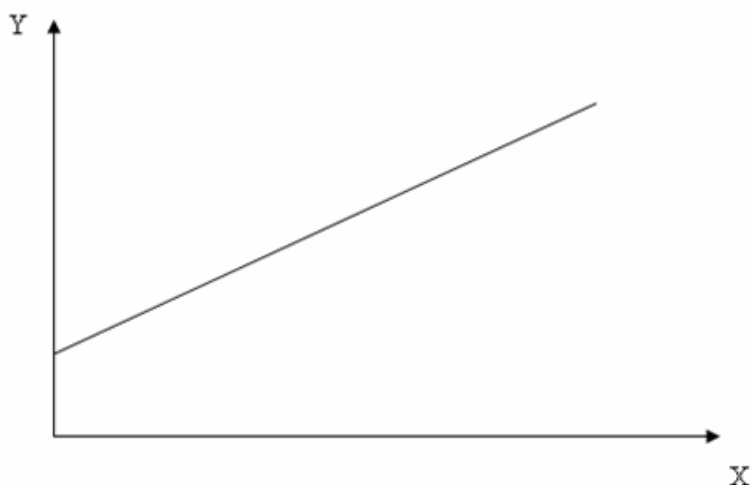
Phần này liệt kê một số dạng hàm được sử dụng phổ biến nhất, cho biết chúng biểu hiện như thế nào, mô tả các tính chất của chúng, và cho chúng ta một số ý tưởng về cách chọn lựa giữa các dạng hàm này.

#### 1.4.1. Dạng Hàm Tuyến tính.

Dạng hàm này có phương trình:

$$Y = \beta_0 + \beta_1 X + \varepsilon.$$

Dạng hàm tuyến tính có thể được mô tả ở dạng như sau:



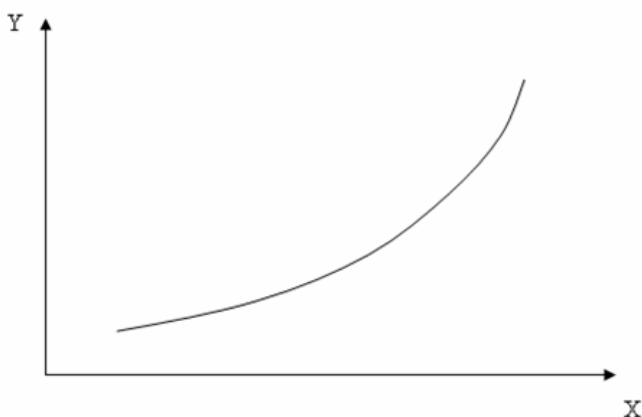
Ưu điểm của dạng hàm tuyến tính là tính đơn giản của nó. Mỗi lần  $X$  tăng thêm một đơn vị thì  $Y$  tăng thêm  $\beta_1$  đơn vị, và điều này đúng bất kể các giá trị của  $X$  và  $Y$  là bao nhiêu. Nhược điểm của dạng hàm tuyến tính cũng chính là tính đơn giản của nó, bất cứ lúc nào tác động của  $X$  phụ thuộc vào các giá trị của  $X$  hoặc  $Y$ , thì dạng hàm tuyến tính không thể là dạng hàm phù hợp. Thí dụ, nếu ta có đường biểu diễn chi phí có dạng  $C = \beta_0 + \beta_1 Q + \varepsilon$ , thì dạng hàm tuyến tính ám chỉ là khi  $Q$  tăng thêm một đơn vị thì chi phí  $C$  tăng thêm  $\beta_1$  đơn vị. Điều này chỉ có thể đúng trong trường hợp chi phí biên không đổi; nó không thể đúng trong trường hợp chi phí biên tăng dần (hay giảm dần). Nếu chúng ta nghĩ rằng chi phí biên tăng dần, chúng ta sẽ không muốn sử dụng dạng hàm tuyến tính.

#### 1.4.2. Dạng Hàm Bậc hai.

Dạng hàm này cho phép giải thích tác động của  $X$  lên  $Y$  phụ thuộc vào giá trị hiện hành của  $X$ . Nó có phương trình:

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \varepsilon.$$

Dạng hàm bậc hai có thể được mô tả ở dạng như sau:



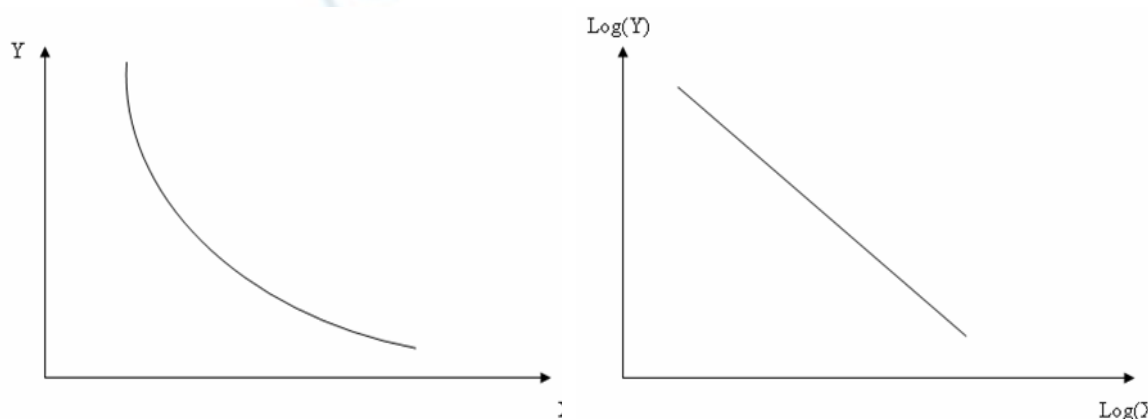
Ưu điểm của dạng hàm bậc hai là khi  $X$  tăng thêm một đơn vị thì  $Y$  tăng thêm  $\beta_1 + 2\beta_2 X$  đơn vị. (Dễ dàng thấy được điều này bằng cách tính  $dY/dX$  từ phương trình nói trên.) Nếu  $\beta_2 > 0$ , thì khi  $X$  tăng lên, tác động bổ sung của  $X$  đến  $Y$  cũng tăng lên; nếu  $\beta_2 < 0$ , thì khi  $X$  tăng lên, tác động bổ sung của  $X$  đến  $Y$  giảm xuống. Nếu ta có đường biểu diễn chi phí  $C = \beta_0 + \beta_1 Q + \beta_2 Q^2 + \varepsilon$ , thì chi phí biên (tức là số đơn vị mà  $C$  tăng lên khi  $Q$  tăng lên thêm một đơn vị) sẽ là  $MC = \beta_1 + 2\beta_2 Q$ . Nếu  $\beta_2 > 0$ , thì ta có chi phí biên tăng dần; nếu  $\beta_2 < 0$ , thì ta có chi phí biên giảm dần. Lý thuyết kinh tế gợi ý rằng chúng ta thường có chi phí biên tăng dần (hay có thể không đổi); chúng ta có thể kiểm định liệu có phải  $\beta_2 > 0$  (hay  $\beta_2 = 0$ ) để kiểm định lý thuyết này.

### 1.4.3. Dạng Hàm Logarít.

Dạng hàm này có phương trình:

$$\log Y = \beta_0 + \beta_1 \log X + \varepsilon.$$

Đồ thị của dạng hàm này có thể được mô tả ở dạng như sau:



Có hai cách để nghĩ về dạng hàm này. Một cách để giải thích dạng hàm này là nếu  $X$  thay đổi 1% thì  $Y$  sẽ thay đổi  $\beta_1\%$ ; đây là tính chất đặc biệt của quan hệ lôgarít. Cách giải thích thứ hai về dạng hàm này là  $\beta_1$  là độ co giãn của  $Y$  theo  $X$ ; điều này suy ra từ định nghĩa của độ co giãn (chúng ta dễ dàng chứng minh điều này bằng một số biến đổi, bắt đầu từ  $d(\log Y)/d(\log X)$  bằng với  $(dY/dX)(X/Y)$  và sắp xếp các số hạng lại). Dạng hàm này thường được sử dụng khi chúng ta quan tâm đến việc ước lượng một loại độ co giãn nào đó. Người ta cũng sử dụng dạng hàm này phổ biến khi chúng ta sử dụng hàm Cobb Douglas; hàm Cobb-Douglas có dạng  $Y = AX^{\beta_1}e^{\varepsilon}$ .

và nếu lấy log ở cả hai vế, chúng ta được:  $\ln Y = \ln(A) + \beta_1 \ln X + \varepsilon$

(Trong Eviews hàm này được viết dưới dạng  $\log Y = \log(A) + \beta_1 \log X + \varepsilon$ )

và ta có thể đặt  $\beta_0 = \log(A)$ . Vì thế cho nên dạng hàm lôgarít thường được sử dụng cho các hàm chi phí, các hàm sản xuất, các hàm hữu dụng, và các hàm khác mà chúng thường được mô tả dưới dạng hàm Cobb-Douglas.

Ví dụ về Hàm sản xuất Cobb-Douglas

$$Y = \beta_1 K^{\beta_2} L^{\beta_3} e^{\varepsilon}$$

Y = sản lượng

K = nhập lượng vốn

L = nhập lượng lao động

Đây là mối quan hệ phi tuyến, nhưng chúng ta có thể biến đổi quan hệ này:

$$\text{Nhu thế: } \ln Y = \ln \beta_1 + \beta_2 \ln K + \beta_3 \ln L + \varepsilon$$

Đây là mô hình tuyến tính trong các tham số nhưng không tuyến tính trong các biến số. Mô hình này tuyến tính theo lôgarít của các biến số. Mô hình này được gọi là mô hình lôgarít-lôgarít, lôgarít kép hay tuyến tính lôgarít

Hệ số độ dốc của một mô hình tuyến tính lôgarít đo lường độ co giãn của Y theo X. Như thế, hệ số nói trên là độ co giãn. Độ co giãn này không đổi đối với X&Y

- +  $\beta_2 + \beta_3$  đo lường hiệu quả theo qui mô. Đáp ứng của sản lượng đối với thay đổi tương xứng trong các nhập lượng.
- + Nếu  $\beta_2 + \beta_3 = 1$ : hiệu quả không đổi. Tăng gấp đôi nhập lượng thì sản lượng sẽ tăng gấp đôi.
- + Nếu  $\beta_2 + \beta_3 < 1$ : hiệu quả giảm dần
- + Nếu  $\beta_2 + \beta_3 > 1$ : hiệu quả tăng dần

Ví dụ khi hồi qui theo dạng hàm này với dữ liệu về nông nghiệp của Đài Loan từ 16 quan sát, người ta thu được kết quả như sau:

$$\ln Y = -3.34 + 0.49 \ln K + 1.50 \ln L$$

$$t \quad (-1.36) \quad (4.80) \quad (0.54)$$

$$R^2 = 0.89$$

Y GNP tính bằng triệu đô la

K là vốn thực tính bằng triệu đô la

L tính bằng triệu ngày công lao động

Độ co giãn của sản lượng theo vốn là 0,49. Giữ nhập lượng lao động không đổi, gia tăng 1% nhập lượng vốn dẫn đến gia tăng 0,49% sản lượng.

Độ co giãn của sản lượng theo lao động là 1,50. Giữ nhập lượng vốn không đổi, gia tăng 1% nhập lượng lao động dẫn đến gia tăng 1,5% sản lượng.

Hiệu quả tăng theo qui mô bởi vì  $\beta_2 + \beta_3 = 1,99$ .

$R^2$  có nghĩa là 89% biến thiên trong lôgarít của sản lượng được giải thích bởi lôgarít của lao động và vốn.

Hay ví dụ chúng ta có thể lập mô hình cầu như một mô hình tuyến tính lôgarít và từ đó ước lượng độ co giãn của cầu tiêu dùng cà phê mỗi ngày:

Giả sử kết quả thu được như sau:

$$\ln Q = 0.78 - 0.25 \ln P_{\text{coffee}} + 0.38 \ln P_{\text{tea}}$$

- + Q là mức tiêu dùng cà phê mỗi ngày
- +  $P_{\text{coffee}}$  là giá cà phê mỗi cân Anh
- +  $P_{\text{tea}}$  là giá trà mỗi cân Anh

Độ co giãn theo giá riêng là -0,25.

- + Giữ các yếu tố khác không đổi, nếu giá gia tăng 1% thì lượng cầu sẽ giảm 0,25%.
- + Đây là không co giãn - giá trị tuyệt đối nhỏ hơn 1.

Độ co giãn theo giá-chéo là 0,38.

- + Giữ các yếu tố khác không đổi, nếu giá trà gia tăng 1%, thì lượng cầu cà phê sẽ gia tăng 0,38%
- + Nếu độ co giãn theo giá-chéo dương, thì cà phê và trà là các sản phẩm thay thế.
- + Nếu độ co giãn theo giá-chéo âm, thì đó là các sản phẩm bổ trợ.

#### **1.4.4. Dạng Hàm Translog.**

Phương trình của dạng hàm này là:

$$\log Y = \beta_0 + \beta_1 \log X + \beta_2 (\log X)^2 + \varepsilon.$$

Dạng hàm này có mối quan hệ với dạng hàm lôgarít giống như mối quan hệ giữa hàm bậc hai và hàm tuyến tính; nó đưa thêm một số hạng bình phương vào phương trình. Trong dạng hàm này, độ co giãn của Y theo X là  $\beta_1 + 2\beta_2 \log X$ .

#### **1.4.5. Dạng Hàm bán-lôgarít (Semilog).**

Dạng hàm này có phương trình sau:

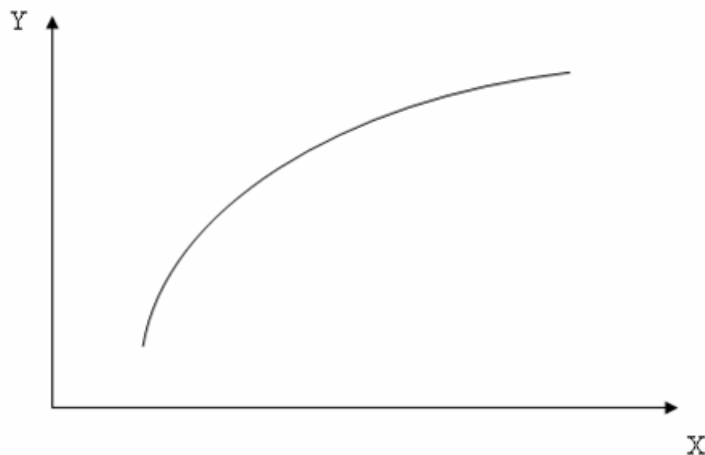
$$\log Y = \beta_0 + \beta_1 X + \varepsilon.$$

Sử dụng mô hình này khi chúng ta quan tâm đến tốc độ tăng trưởng của biến nào đó như GNP hay mức việc làm.

Dạng hàm bán-lôgarít có tính chất là nếu  $X$  tăng thêm 1 đơn vị thì  $Y$  tăng thêm  $[\beta_1 \cdot 100] \%$ . Đây không phải là một tính chất được mong muốn một cách phổ biến, nhưng có một số ứng dụng hữu ích cho dạng hàm này.

Ví dụ, quan hệ giữa tiền lương và trình độ giáo dục hầu như luôn luôn được biểu hiện dưới dạng hàm này như là  $\log(SAL) = \beta_0 + \beta_1 ED + \varepsilon$ . Điều này có nghĩa là nếu trình độ giáo dục của một người tăng 1 năm thì tiền lương của người đó tăng  $[\beta_1 \cdot 100] \%$ . Thí dụ, nếu  $\beta_1 = 0,08$ , nó có nghĩa là một năm tăng thêm trong trình độ giáo dục làm tăng tiền lương thêm 8%. Khi  $X$  tăng lên thì độ dốc của đường biểu diễn sẽ trở nên rất lớn, bởi vì khi  $X$  tăng lên thì tỷ lệ phần trăm gia tăng của  $X$  cũng lớn hơn.

Chúng ta cũng có thể đặt lôgarít cho  $X$ , nghĩa là dạng hàm trên trở thành,  $Y = \beta_0 + \beta_1 \cdot \log X + \varepsilon$  và điều này có nghĩa là khi  $X$  tăng 1%, thì  $Y$  tăng  $[\beta_1/100]$  đơn vị. Trong trường hợp này, khi  $X$  càng lớn thì độ dốc càng nhỏ, bởi vì  $X$  cần gia tăng nhiều hơn mới tạo ra được 1% gia tăng. Dạng hàm này không được sử dụng rộng rãi; dưới đây là đồ thị của nó.

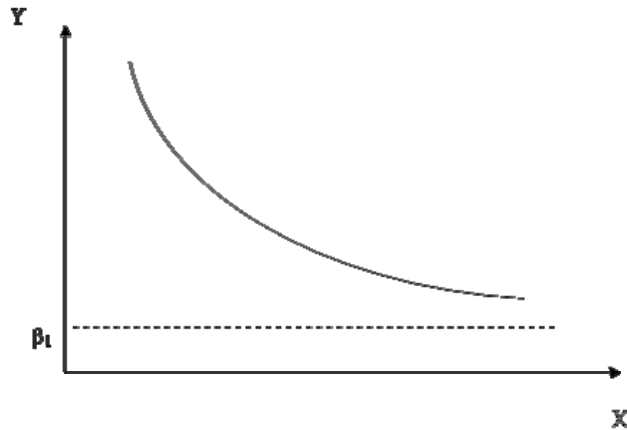


#### 1.4.6. Dạng Hàm Nghịch đảo.

Dạng hàm này có phương trình như sau:

$$Y = \beta_0 + \beta_1 \left( \frac{1}{X} \right) + \varepsilon.$$

Đồ thị của dạng hàm này có thể được mô tả ở dạng như sau:



Dạng hàm nghịch đảo thường được sử dụng khi  $Y$  và  $X$  đều dương và khi đường biểu diễn quan hệ giữa chúng có lẽ dốc xuống (nghĩa là,  $\beta_0 > 0$  và  $\beta_1 > 0$ ). Trong trường hợp này, dạng hàm tuyến tính không được tốt bởi vì đường biểu diễn sẽ cắt trục tọa độ và  $Y$  sẽ trở nên âm đối với các giá trị  $X$  đủ lớn. Dạng hàm này thường được sử dụng cho các đường (cong) như đường cầu hay đường chi phí cố định (chi phí cố định trung bình trong sản xuất giảm xuống liên tục khi sản lượng tăng lên) cần có tính chất này.

Nếu chúng ta có nhiều biến độc lập, thì dạng quan hệ giữa  $Y$  và mỗi biến có thể giống nhau hoặc có thể khác nhau. Chúng ta có thể sử dụng cùng một dạng hàm như nhau cho mỗi biến; ví dụ, nếu chúng ta nghĩ rằng sản lượng phụ thuộc vào ba nhập lượng khác nhau thì có thể sử dụng dạng hàm tuyến tính cho mỗi nhập lượng:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon.$$

Hay dạng hàm translog cho mỗi nhập lượng:

$$\log Y = \beta_0 + \beta_1 \log X_1 + \beta_2 (\log X_1)^2 + \beta_3 \log X_2 + \beta_4 (\log X_2)^2 + \varepsilon.$$

Hay tổng quát hơn:

$$\log Y = \beta_0 + \sum_{i=1}^k \beta_i X_i + \sum_{i=1}^k \sum_{j=1}^k \gamma_{ij} (\log X_i)(\log X_j) + \varepsilon.$$

Chúng ta cũng có thể kết hợp vài dạng hàm khác nhau trong một hồi qui, thí dụ:

$$Y = \beta_0 + \beta_1 \log X + \beta_2 X_2 + \beta_3 X_2^2 + \beta_4 \left(\frac{1}{X_3}\right) + \varepsilon.$$

mặc dù, nếu làm thế, chúng ta thường phải có các lý do thỏa đáng để nghĩ rằng hình dạng của quan hệ giữa  $Y$  và  $X_1$  là khác với các hình dạng của quan hệ giữa  $Y$  và  $X_2$ , và  $Y$  và  $X_3$ .

### 1.5. Các loại dữ liệu cho phân tích kinh tế lượng

Để ước lượng mô hình kinh tế đã đưa ra, cần có mẫu dữ liệu về các biến phụ thuộc và biến độc lập.

Có 3 loại số liệu được sử dụng để phân tích:

1. Các số liệu theo thời gian (chuỗi thời gian)
2. Các số liệu chéo
3. Các số liệu hỗn hợp của hai loại trên.

Các số liệu có thể dạng số lượng (như. GDP, tỷ giá hối đoái, Giá chứng khoán), hay dạng chất lượng (như. Nam/ nữ; có gia đình/ chưa có gia đình; Quá trình sản xuất A/quá trình sản xuất B).

#### 1.5.1. *Số liệu theo thời gian:*

Là số liệu được thu thập trong một thời kỳ, như:

- Quan sát mức lạm phát và thất nghiệp của Mỹ từ 1962-1995
- Quan sát GDP của Mỹ từ 1960-1992
- Quan sát khả năng sinh lời của một công ty trong hơn 20 năm
- Quan sát giá vàng hàng ngày lúc đóng cửa trong hơn 30 năm.

Ví dụ, giả sử một thành phố muốn dự báo nhu cầu nhà ở cho năm hoặc mười năm trong tương lai. Việc này đòi hỏi phải xác định các biến có ảnh hưởng đến nhu cầu nhà ở của thành phố đó trong quá khứ, có được chuỗi dữ liệu theo thời gian trong nhiều năm ở quá khứ, và sử dụng chúng vào một mô hình thích hợp để tạo các giá trị dự báo của nhu cầu tương lai. Khoảng thời gian hoặc **thời đoạn của chuỗi thời gian** sẽ là hàng năm, hàng quý hoặc hàng tháng, tùy theo thành phố đó muốn xem xét thay đổi trong nhu cầu nhà ở hàng năm, hàng quý hay hàng tháng. Loại dữ liệu sẵn có thường sẽ quyết định thời đoạn của dữ liệu thu thập.

#### 1.5.2. *Số liệu chéo:*

Là số liệu về một hay nhiều biến được thu thập tại một thời điểm ở nhiều địa phương, đơn vị khác nhau.

- Quan sát chiều cao và cân nặng của 1000 người
- Quan sát thu nhập, trình độ học vấn, và cân nặng của 1000 người
- Quan sát khả năng sinh lời của 20 công ty
- Quan sát GDP trên đầu người, dân số, và chi phí quốc phòng thực tế của 80 quốc gia.

Ví dụ, khi chúng ta muốn xem xét thu nhập ảnh hưởng như thế nào đến tiêu dùng của một người. Việc này đòi hỏi phải quan sát thu nhập và tiêu dùng của nhiều người trong một khoảng thời gian xác định.

### **1.5.3. Số liệu hỗn hợp:**

Số liệu hỗn hợp theo không gian và thời gian, ví dụ như:

- Quan sát tỷ lệ lạm phát và mức tăng trưởng của 15 quốc gia trong khoảng thời gian từ 1970-1995
- Quan sát mức sản lượng và mức giá của 100 ngành trong hơn 12 quý
- Quan sát khả năng sinh lời của 20 công ty trong hơn 20 năm

Số liệu chuỗi thời gian thường người ta ký hiệu là  $t$  và tổng số quan sát là  $T$ , còn đối với số liệu chéo ta ký hiệu quan sát là  $i$  và tổng số quan sát là  $N$ .

Dữ liệu có thể thu thập được tại những nguồn sẵn có. Tuy nhiên, trong nhiều trường hợp những nguồn này không đủ để giải quyết vấn đề đặt ra hoặc những dữ liệu này không có sẵn. Trong trường hợp như vậy, cần tiến hành những khảo sát để thu thập các thông tin cần thiết. Ví dụ, chúng ta muốn quan tâm đến việc nghiên cứu xem người tiêu dùng sẽ phản ứng như thế nào đối với chính sách giá điện. Chính sách giá điện trong ngày là giá điện sẽ thay đổi theo những giờ khác nhau trong ngày, với giá cao trong những giờ cao điểm và giá thấp trong những giờ thấp điểm. Để có được dữ liệu phù hợp người ta chọn một số khách hàng và lắp đặt đồng hồ để ghi lại lượng điện sử dụng từng giờ trong ngày. Lượng điện tiêu thụ được thu thập trong vòng một năm, như vậy có được dữ liệu theo chuỗi thời gian cho một nhóm các hộ tiêu thụ nào đó.

## **TÓM TẮT**

Kinh tế lượng liên quan đến ước lượng các mối liên hệ kinh tế, kiểm định giả thuyết các lý thuyết kinh tế, và dự báo các biến kinh tế hoặc các biến số khác. Khi nghiên cứu, chúng ta thường phải bắt đầu với một tập hợp các lý thuyết kinh tế, sau đó kết hợp chúng với những nhận định trực giác (hoặc kinh nghiệm, nghiên cứu trong quá khứ) để xây dựng một mô hình kinh tế lượng. Quá trình này liên quan đến quyết định chọn một hay nhiều biến phụ thuộc và xác định các biến độc lập có ảnh hưởng đến các biến phụ thuộc.

Bước tiếp theo là thu thập dữ liệu tương ứng. Khi có được các dữ liệu này, chúng ta sẽ ước lượng các thông số của một hoặc nhiều mô hình sơ bộ. Các mô hình này sẽ được kiểm định nhiều lần, dựa vào những kiểm định này, các mô hình được thiết lập lại và ước lượng lại cho đến khi thỏa mãn. Mô hình cuối cùng có thể được dùng để xây dựng các chính sách hoặc để dự báo các giá trị của các biến phụ thuộc trong nhiều tình huống khác nhau.

## CHƯƠNG II

### MÔ HÌNH HỒI QUY HAI BIẾN, ƯỚC LƯỢNG VÀ KIỂM ĐỊNH

Ở chương I phát biểu rằng bước đầu tiên trong phân tích kinh tế lượng là việc thiết lập mô hình mô tả được hành vi của các đại lượng kinh tế. Tiếp theo đó nhà phân tích kinh tế/ kinh doanh sẽ thu thập những dữ liệu thích hợp và ước lượng mô hình nhằm hỗ trợ cho việc ra quyết định. Trong chương này sẽ giới thiệu mô hình đơn giản nhất và phát triển các phương pháp ước lượng, phương pháp kiểm định giả thuyết và phương pháp dự báo. Mô hình này đề cập đến biến độc lập ( $Y$ ) và một biến phụ thuộc ( $X$ ). Đó chính là mô hình hồi quy tuyến tính hai biến (thường gọi là mô hình hồi quy đơn). Mặc dù đây là một mô hình đơn giản, và đôi khi có thể là phi thực tế, nhưng việc hiểu biết những vấn đề cơ bản trong mô hình này là nền tảng cho việc tìm hiểu những mô hình phức tạp hơn. Thực tế, mô hình hồi quy đơn tuyến tính có thể giải thích cho nhiều phương pháp kinh tế lượng. Trong chương này chỉ đưa ra những kết luận căn bản về mô hình hồi quy tuyến tính đơn biến.

Mục tiêu đầu tiên của một nhà kinh tế lượng là làm sao sử dụng dữ liệu thu thập được để ước lượng hàm hồi quy của tổng thể, đó là, ước lượng tham số của tổng thể  $\beta_1$  và  $\beta_2$ . Ký hiệu  $\hat{\beta}_1$  là ước lượng mẫu của  $\beta_1$  và  $\hat{\beta}_2$  là ước lượng mẫu của  $\beta_2$ . Khi đó mối quan hệ trung bình ước lượng là  $\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X$ . Đây được gọi là hàm hồi quy của mẫu.

Thuật ngữ *đơn* trong mô hình hồi quy tuyến tính đơn được sử dụng để chỉ rằng chỉ có duy nhất một biến giải thích ( $X$ ) được sử dụng trong mô hình.

Trong chương tiếp theo khi nói về *mô hình hồi quy đa biến* sẽ bổ sung thêm nhiều biến giải thích khác. Thuật ngữ *hồi quy* xuất phát từ Francis Galton (1886), người đặt ra mối liên hệ giữa chiều cao của người con trai với chiều cao của người cha và quan sát thực nghiệm cho thấy có một xu hướng giữa chiều cao trung bình của người con trai với chiều cao của những người cha của họ để “hồi quy” cho chiều cao trung bình của toàn bộ tổng thể.  $\beta_1 + \beta_2 X_i$  gọi là *phần xác định* của mô hình và là trung bình có điều kiện của  $Y$  theo  $X$ , đó là  $E(Y_i) = \beta_1 + \beta_2 X_i$ . Thuật ngữ *tuyến tính* dùng để chỉ rằng bản chất của các *thông số của tổng thể*  $\beta_1$  và  $\beta_2$  là tuyến tính (bậc nhất).

#### 2.1. Khái niệm hàm hồi quy tổng thể

Tổng thể là toàn bộ các quan sát về các đối tượng hay con người cho mục đích nghiên cứu. Mục tiêu đầu tiên của một nhà kinh tế lượng là làm sao sử dụng dữ liệu thu thập được để ước lượng hàm hồi quy của tổng thể, đó là, ước lượng tham số của tổng thể  $\beta_1$  và  $\beta_2$ .

Cho  $Y$  là biến được giải thích, chọn  $X_2, X_3, \dots, X_k$  là biến giải thích.  $Y$  là ngẫu nhiên và có 1 phân phối xác suất nào đó.

$\Rightarrow$  tồn tại  $E(Y|X_2, X_3, \dots, X_k) =$  giá trị xác định

Do vậy  $F(X_2, X_3, \dots, X_k) = E(Y|X_2, X_3, \dots, X_k)$  là hàm hồi quy tổng thể của  $Y$  theo  $X_2, X_3, \dots, X_k$  (PRF-population regression function), hàm phụ thuộc ở mức độ trung bình của  $Y$  theo  $X$ .

Với một cá thể  $i$ , tồn tại  $(X_{2i}, X_{3i}, \dots, X_{ki}, Y_i)$

Ta có  $Y_i \neq F(X_2, X_3, \dots, X_k) \Rightarrow u_i = Y_i - F$

Do vậy:  $Y_i = E(Y|X_2, X_3, \dots, X_k) + u_i$

Hồi qui tổng thể PRF:

$$Y = E(Y|X) + U$$

$$E(Y|X) = F(X)$$

## 2.2. Hàm hồi qui mẫu

Do không biết tổng thể, nên chúng ta không biết giá trị trung bình tổng thể của biến phụ thuộc là đúng ở mức độ nào. Do vậy chúng ta phải dựa vào dữ liệu mẫu để ước lượng.

Trên thực tế khi tổng thể lớn, tồn tại  $F$  nhưng không tìm được chính xác do:

- ✓ Không quan sát được (do thời gian hay tài chính không cho phép..)
- ✓ Tổng thể biến động
- ✓ Đặc điểm thông tin: không cần quan sát

Do vậy người ta phải tiến hành chọn mẫu, mẫu là một nhóm hay một bộ phận của tổng thể.

Hồi qui mẫu:

$$\text{Cho PRF: } Y = F(x_2, x_3, \dots, x_k) + u$$

Trên một bộ phận (mẫu) có  $n$  cá thể gọi  $\hat{Y} = \hat{F}(X_2, X_3, \dots, X_k)$  là hồi qui mẫu (SRF - Sample regression function)

Với một cá thể mẫu  $Y_i \neq \hat{F}(X_{2i}, X_{3i}, \dots, X_{ki})$

Sinh ra  $e_i = Y_i - \hat{F}(X_{2i}, X_{3i}, \dots, X_{ki}) = Y_i - \hat{Y}_i$ ;  $e_i$  gọi là *Phần dư SRF*

Ký hiệu  $\hat{\beta}_1$  là ước lượng mẫu của  $\beta_1$  và  $\hat{\beta}_2$  là ước lượng mẫu của  $\beta_2$ . Khi đó mối quan hệ trung bình ước lượng là  $\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X$ . Đây được gọi là hàm hồi quy của mẫu.

Ước lượng SRF: Chọn 1 phương pháp nào đó để ước lượng các tham số của  $F$  qua việc tìm các tham số  $\hat{F}$  và lấy giá trị quan sát của các tham số này làm giá trị xấp xỉ cho tham số của  $F$ .

## 2.3. Phương pháp bình phương nhỏ nhất

### 2.3.1. Tư tưởng của phương pháp bình phương nhỏ nhất

Trong kinh tế lượng, thủ tục ước lượng được dùng phổ biến nhất là phương pháp bình phương nhỏ nhất. Tiêu chuẩn tối ưu được sử dụng bởi phương pháp bình phương nhỏ nhất là cực tiểu hóa hàm mục tiêu.

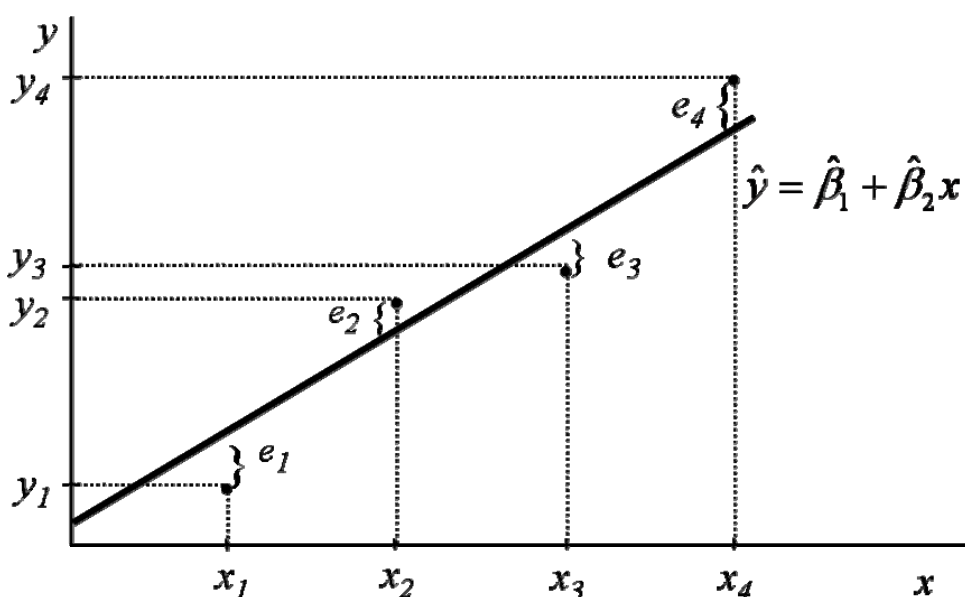
Phương pháp bình phương nhỏ nhất là một phương pháp được đưa ra bởi nhà toán học Đức Carl Friedrich Gauss, đây là một phương pháp mạnh và được rất nhiều người sử dụng, nó thường được ký hiệu là OLS (ordinary least squares). Tư tưởng của phương pháp này là cực tiểu tổng bình phương các phần dư. Do đó có thể nói để có được đường hồi qui “thích hợp” nhất, chúng ta chọn các ước lượng của tung độ gốc và độ dốc sao cho phần dư là nhỏ.

Chúng ta đặt:

$y_i$  ký hiệu giá trị thực của biến  $y$  tại quan sát  $i$

$\hat{y}_i$  ký hiệu giá trị của hàm hồi qui mẫu

$e_i$  ký hiệu phần dư,  $y_i - \hat{y}_i$



Do đó cực tiểu hoá  $\sum (y_i - \hat{y}_i)^2$  sẽ tương đương với cực tiểu  $\sum e_i^2$  từ đó tìm ra  $\beta_1$  và  $\beta_2$ .

Chúng ta có thể mô tả tổng quát như sau:

Xét hàm hồi qui tổng thể (PRF):  $Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k + u$