

Econometrics

Kinh tế lượng

Sách tham khảo:

Bài giảng Kinh tế lượng – PGS.TS Nguyễn Quang Đông -
Đại học KTQD Hà Nội

Kết cấu học phần: gồm 3 tín chỉ

- Chương 1: Mở đầu
- Chương 2: Mô hình hồi quy đơn
- Chương 3: Mô hình hồi quy bội
- Chương 4: Hồi quy với biến giả
- Chương 5: Đa cộng tuyến
- Chương 6: Phương sai của sai số thay đổi
- Chương 7: Tự tương quan
- Chương 8: Chọn mô hình và kiểm định dạng mô hình

Chương 1: Mở đầu

I. Khái quát về Kinh tế lượng

1. Kinh tế lượng là gì?

- “Kinh tế lượng”: Econometrics – “đo lường kinh tế” do Gs Ragnar Frisch đưa ra năm 1930
 - 1936: Tinbergen trình bày mô hình kinh tế lượng đầu tiên
 - 1950: Lawrence Klein đưa ra một số mô hình kinh tế cho Mỹ. Hiện nay, ông phụ trách dự án dự báo kinh tế thế giới cho LHQ
- ⇒ Kinh tế lượng là môn khoa học đo lường các mối quan hệ kinh tế diễn ra trong thực tế, là kết hợp giữa lý thuyết kinh tế, thống kê toán học và phần mềm máy tính.
- ⇒ Là một môn khoa học độc lập

2. Phương pháp luận kinh tế lượng:

- Bước 1: Nêu các giả thiết về mối quan hệ kinh tế cần phân tích.

⇒ Vd: *Mức tiêu dùng của các hộ gia đình quan hệ tỷ lệ thuận với thu nhập khả dụng của họ.*

- Bước 2: Thiết lập các mô hình toán học mô tả mối quan hệ này (xác định dạng hàm...)

⇒ Vd: Mô hình giá nhà:

$$Gi_s = \alpha + \beta Di0ntfch$$

$$Gi_s = \beta 1 +$$

$$\beta 2 Di0ntfch + \beta 3 Kichthucvuon + \beta 4 SèphBngt\%m + \beta 5 SèphBngngñ.$$

⇒ Vd: Mô hình miêu tả quan hệ giữa việc hút thuốc và tử vong do ung thư phổi

$$DEATHS = \alpha + \beta SMOKING$$

- Bước 3: Thu thập số liệu => đòi hỏi kích thước mẫu khá lớn
- Bước 4: Ước lượng các tham số của mô hình => phương pháp ước lượng thông thường là phương pháp OLS
- Bước 5: Phân tích kết quả => đánh giá kết quả nhận được có phù hợp lý thuyết kinh tế hay không?
- Bước 6: Dự báo => dự báo trong các chu kỳ kế tiếp gồm dự báo giá trị trung bình và dự báo giá trị cá biệt.
- Bước 7: Sử dụng mô hình để đề ra các chính sách kinh tế.

Sơ đồ các bước

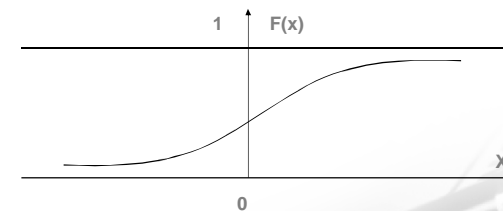


II. Cơ sở thống kê toán của Kinh tế lượng:

1. Một số khái niệm cơ bản:

- **Xác suất:**
Là khả năng xuất hiện của 1 sự kiện
Vd: Một phép thử liên quan đến sự kiện A, trong n phép thử, A xuất hiện m lần thì $P(A) = m/n$
- **Biến ngẫu nhiên:**
Là biến nhận các giá trị một cách ngẫu nhiên
=> Biến ngẫu nhiên rời rạc
=> Biến ngẫu nhiên liên tục
- **Hàm phân phối xác suất:**
Hàm PPXS của biến ngẫu nhiên X là xác suất của sự kiện $X < x$. $F(x) = P(X < x)$

=> Nếu X là biến ngẫu nhiên liên tục thì F(x) có dạng:



Hàm phân phối xác suất là hàm không giảm nghĩa là, nếu $x_2 > x_1$ thì $F(x_2) > F(x_1)$.

▪ Hàm mật độ phân phối xác suất:

Xác định xác suất để một biến ngẫu nhiên liên tục nhận giá trị tại một điểm: $f(x) = P(X = x)$

Cách xác định:

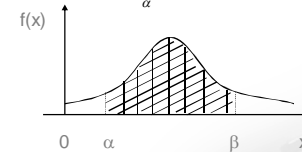
- ✓ Tính xác suất để biến NN nhận giá trị trong một khoảng $P(x < X < x + \Delta x)$
- ✓ Tính xác suất trung bình trên đoạn Δx : $\lim(P(x < X < x + \Delta x)/\Delta x = f(x))$ khi $\Delta x \rightarrow 0$
Ta thấy rằng: $P(x < X < x + \Delta x) = P(X < x + \Delta x) - P(X < x) = F(x + \Delta x) - F(x) = \Delta F(x)$.
Như vậy: $f(x) = \lim \Delta F(x)/\Delta x = F'(x)$.

Các tính chất của $f(x)$:

TC1: $f(x) \geq 0$.

TC2: $F(x) = F(x) - F(-\infty) = \int_{-\infty}^x f(x)dx$

TC3: $P(\alpha < x < \beta) = \int_{\alpha}^{\beta} f(x)dx = F(\beta) - F(\alpha) = P(x < \beta) - P(x < \alpha)$



TC4: $\int_{-\infty}^{+\infty} f(x)dx = 1$

▪ Kỳ vọng toán học của biến ngẫu nhiên:

⇒ Đối với biến rời rạc: $E(X) = \sum X_i p_i$

⇒ Đối với biến liên tục:

$E(X) = \int_a^b xf(x)dx$ hữu hạn trong đoạn $[a,b]$; $= \int_{-\infty}^{+\infty} xf(x)dx$ vô hạn.

VD: Gọi X là số chấm xuất hiện khi gieo một con xúc sắc. Tính kỳ vọng $E(X)$.

X	1	2	3	4	5	6
P(X)	1/6	1/6	1/6	1/6	1/6	1/6

⇒ $E(X) = 1.(1/6) + 2.(1/6) + 3.(1/6) + 4.(1/6) + 5.(1/6) + 6.(1/6) = 3.5$

VD: Một người mua 10.000 đồng xổ số lô tô 2 số với luật chơi như sau: thắng 700000 đồng nếu số mua trùng với 2 số cuối của giải độc đắc và không được đồng nào nếu không trùng. Hãy tìm số tiền thắng trung bình của một lần chơi?

⇒ Gọi X là số tiền thắng của một lần chơi, rõ ràng X nhận các giá trị 0đ và 700.000đ với xác suất tương ứng là 99% và 1%. Từ đó số tiền thắng trung bình chính là: $E(X) = 0 \times 99\% + 700000 \times 1\% = 7000đ$

▪ Phương sai:

Biến ngẫu nhiên X có kỳ vọng là $E(X) = m$ thì kỳ vọng của biến $(X-m)^2$ được gọi là phương sai $Var(X)$.

$Var(X) = E(X-m)^2 = E(X^2) - 2E(X).E(m) + E(m^2) = E(X^2) - m^2$

⇒ $X-E(X)$ chính là độ lệch của biến X so với trung bình của nó, do đó phương sai chính là trung bình của bình phương độ lệch đó.

⇒ Phương sai là một số không âm và không cùng đơn vị đo với X

⇒ Sai số tiêu chuẩn: $\sigma(X) = \sqrt{var(X)}$

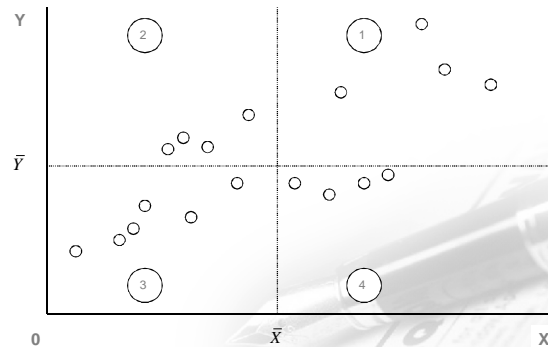
- **Mốt (mode)** là giá trị của biến ngẫu nhiên X có khả năng xuất hiện lớn nhất trong một khoảng nhất định của nó.
- **Trung vị (median)** là giá trị của biến ngẫu nhiên X chia phân phối thành hai phần có xác suất giống nhau: $P(X < \text{med}X) = P(X \geq \text{med}X) = 1/2$.
- **Đồng phương sai (hiệp phương sai) và hệ số tương quan:**
Là các cách để đo lường mức độ quan hệ “chặt” giữa hai biến ngẫu nhiên.

- **Đồng phương sai:**

$$\text{Cov}(X, Y) = E[(X - m_X) \cdot (Y - m_Y)]$$

Ước lượng của thông số này là *phương sai mẫu*:

$$\hat{\sigma}_{XY} = S_{XY} = \frac{1}{n-1} \sum (x_i - \bar{x})(y_i - \bar{y})$$



- **Hệ số tương quan:**

$$= \frac{\text{Cov}(X, Y)}{\sqrt{\text{var}(X) \cdot \text{var}(Y)}}$$

⇒ Hệ số tương quan mẫu:

$$r_{XY} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}}$$

Nếu X và Y có tỷ lệ thuận thì $r > 0$

Nếu X và Y có tỷ lệ nghịch thì $r < 0$

Nếu X và Y không có tương quan thì $r = 0$

2. Các phân phối xác suất thường gặp

2.1. Phân phối chuẩn (Gaoxơ):

- K/n: Biến ngẫu nhiên X có kỳ vọng là m và phương sai σ^2 gọi là phân phối chuẩn và được kí hiệu $N(m, \sigma^2)$ nếu nó có hàm mật độ phân phối xác suất như sau:

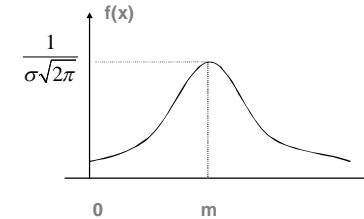
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-m)^2/2\sigma^2}$$

Hoặc hàm phân phối xác suất:

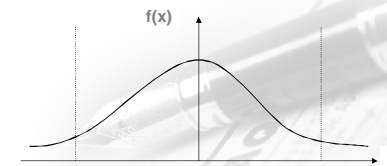
$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(x-m)^2}{2\sigma^2}} dx$$

Kí hiệu: $X \sim N(m, \sigma^2)$.

- Đồ thị:



- Phân phối chuẩn tắc: $N(0, 1)$.

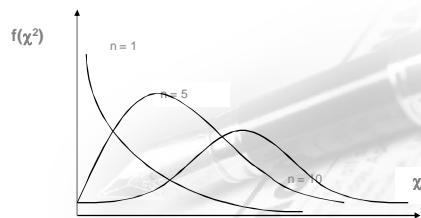


2.2. Phân phối Chi bình phương χ^2

- K/n: Phân phối của tổng bình phương của n biến ngẫu nhiên chuẩn tắc độc lập được gọi là phân phối Chi bình phương với n bậc tự do và được ký hiệu là χ^2 .

Giả sử n biến Z_1, Z_2, \dots, Z_n tuân theo $N(0, 1)$ thì $U = Z_1^2 + Z_2^2 + \dots + Z_n^2$ tuân theo χ^2_n

- Đồ thị:

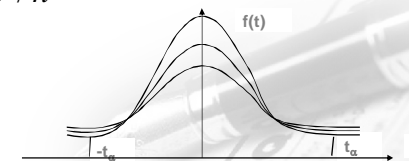


2.3. Phân phối Student t:

- K/n: Phân phối của tỉ lệ giữa một biến chuẩn với căn bậc hai của χ^2_n bậc tự do độc lập được gọi là phân phối t Student với n bậc tự do

$$\Rightarrow t_n = Z / \sqrt{U/n}$$

- Đồ thị:



2.4. Phân phối Fisher F:

- K/n: Phân phối F là tỉ lệ giữa hai biến có phân phối Chi bình phương độc lập

Giả sử $U \sim \chi^2_n$ và $V \sim \chi^2_m$ thì $F = (U/m)/(V/n) \sim F_{m,n}$

- Phân phối F có hình dáng tương tự Chi bình phương

3. Ước lượng:

- K/n: Ước lượng là công thức xác định các thông số cần thiết. Mỗi giá trị bằng số là một giá trị UL
- Phương pháp ước lượng: phương pháp mômen, phương pháp bình phương nhỏ nhất OLS và phương pháp ước lượng hợp lý tối đa ML
- Ước lượng điểm và ước lượng khoảng: Ký hiệu tham số cần UL là θ (tê-ta), ước lượng của nó là $\hat{\theta}$ (tê-ta mũ) là UL điểm.
 $(\hat{\theta}_1, \hat{\theta}_2): P(\hat{\theta}_1 < \theta < \hat{\theta}_2) = 1 - \alpha$ được gọi là UL khoảng

- Các tiêu chuẩn:

- Tính không chệch:

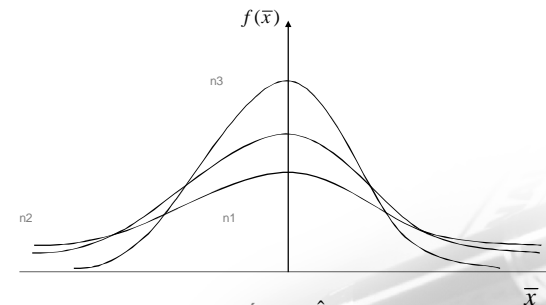
Hàm ước lượng $\hat{\theta}$ của θ gọi là không chệch nếu $E(\hat{\theta}) = \theta$.

- Tính hiệu quả:

Hàm ước lượng $\hat{\theta}$ của θ gọi là hiệu quả nếu $\text{var}(\hat{\theta}) = \min$

- Tính nhất quán (vững):

Hàm ước lượng $\hat{\theta}$ của θ gọi là vững nếu khi n tiến đến vô cùng thì $\hat{\theta}$ hội tụ tới θ .



4. Kiểm định giả thiết

- Giả thiết thống kê: Là một mệnh đề có liên quan tới quy luật xác suất của một đại lượng ngẫu nhiên nào đó.
- ⇒ Vd: Hai biến ngẫu nhiên chuẩn X và Y có kỳ vọng bằng nhau.
- Kiểm định một giả thiết thống kê: là căn cứ vào các mẫu thu được để có quyết định nên bác bỏ hay nên giữ giả thiết đó.
- ⇒ Giả thiết cần kiểm định gọi là giả thiết không và được ký hiệu là H_0 , giả thiết H_1 là giả thiết đối lập.

▪ Quy tắc kiểm định giả thiết:

- Muốn khảo sát một giả thiết liên quan đến quy luật của biến ngẫu nhiên X người ta lấy mẫu x_1, x_2, \dots, x_n
- Dựa vào mẫu xây dựng được trị số k . Với giả thiết H_0 đã cho thì k có một phân phối xác suất nhất định. Ta tìm được k_1, k_2 sao cho $P(k_1 < k < k_2) = 1 - \alpha$.
- Trong thực tế một sự kiện có xác suất nhỏ thì khó xảy ra trong một vài lần quan sát (nguyên lý xác suất nhỏ). Vì vậy nếu một sự kiện có xác suất nhỏ xảy ra thì ta bác bỏ nó.
 - ⇒ k ở ngoài khoảng (k_1, k_2) :
Một sự kiện với xác suất α nhỏ đã xảy ra trong một vài lần quan sát, theo nguyên lý trên ta phải bác bỏ giả thiết H_0 .
Các miền $(-\infty, k_1)$, $(k_2, +\infty)$ là miền bác bỏ giả thiết.
 - ⇒ k ở trong khoảng (k_1, k_2) :
Chấp nhận giả thiết H_0

▪ Các sai lầm kiểm định:

• Sai lầm loại 1:

Giả thiết H_0 là đúng mà ta lại bác bỏ nó. Xác suất của sai lầm này chính là mức ý nghĩa α của kiểm định (xác suất của sự kiện $k < k_1$ và $k > k_2$).

• Sai lầm loại 2:

Giả thiết H_0 là sai mà ta lại chấp nhận nó.