Bài giảng

Phương pháp định lượng trong quản lý

TS. Phạm Cảnh Huy
Khoa Kinh tế và quản lý – ĐHBKHN

Nội dung

- **Mục tiêu học phần:** Phương pháp định lượng trong quản lý giúp cho học viên hiểu và vận dụng được các phương pháp định lượng trong việc ra các quyết định trong quản lý bằng việc ứng dụng những mô hình và các công cụ toán học. Ngoài ra còn cung cấp cho học viên những kỹ năng cần thiết để thực hiện các phân tích định lượng và đánh giá các kết quả từ phân tích định lượng. Thêm nữa môn học còn giúp học viên giải quyết được các bài toán thực tế nhờ công cụ Máy tính để có được một quyết định tốt nhất trong quản lý.
- **Nội dung tóm tắt học phần:** Cung cấp kiến thức cơ bản về phân tích định lượng, ứng dụng phân tích hồi qui trong các nghiên cứu định lượng, cùng những kiến thức cơ bản về lý thuyết toán tối ưu áp dụng trong hoạt động kinh doanh cũng như trong phân tích ra quyết định.

Nội dung

Tài liệu tham khảo:

- Anderson Sweeney Williams, *Study guide for Quantitative methods for business*, Thomson South-Western 2001
- Anderson Sweeney Williams, *An introduction to Management Science, Quantitative Approaches to Decision Making*, Thomson South-Western 2003
- Frederick S.Hillier, *Introduction to Operations Reasearch*, McGraw-Hill 2001
- Damodar N.Gujarati, *Basic Econometrics*, McGraw-Hill 2004
- TS. Phạm Cảnh Huy, *Bài giảng kinh tế lượng*, Nhà xuất bản Đại học Bách khoa Hà Nội 2008
- PGS. TS. Nguyễn Hải Thanh, *Toán ứng dụng (giáo trình sau đại học)*, Nhà xuất bản Đại học sư phạm 2005.

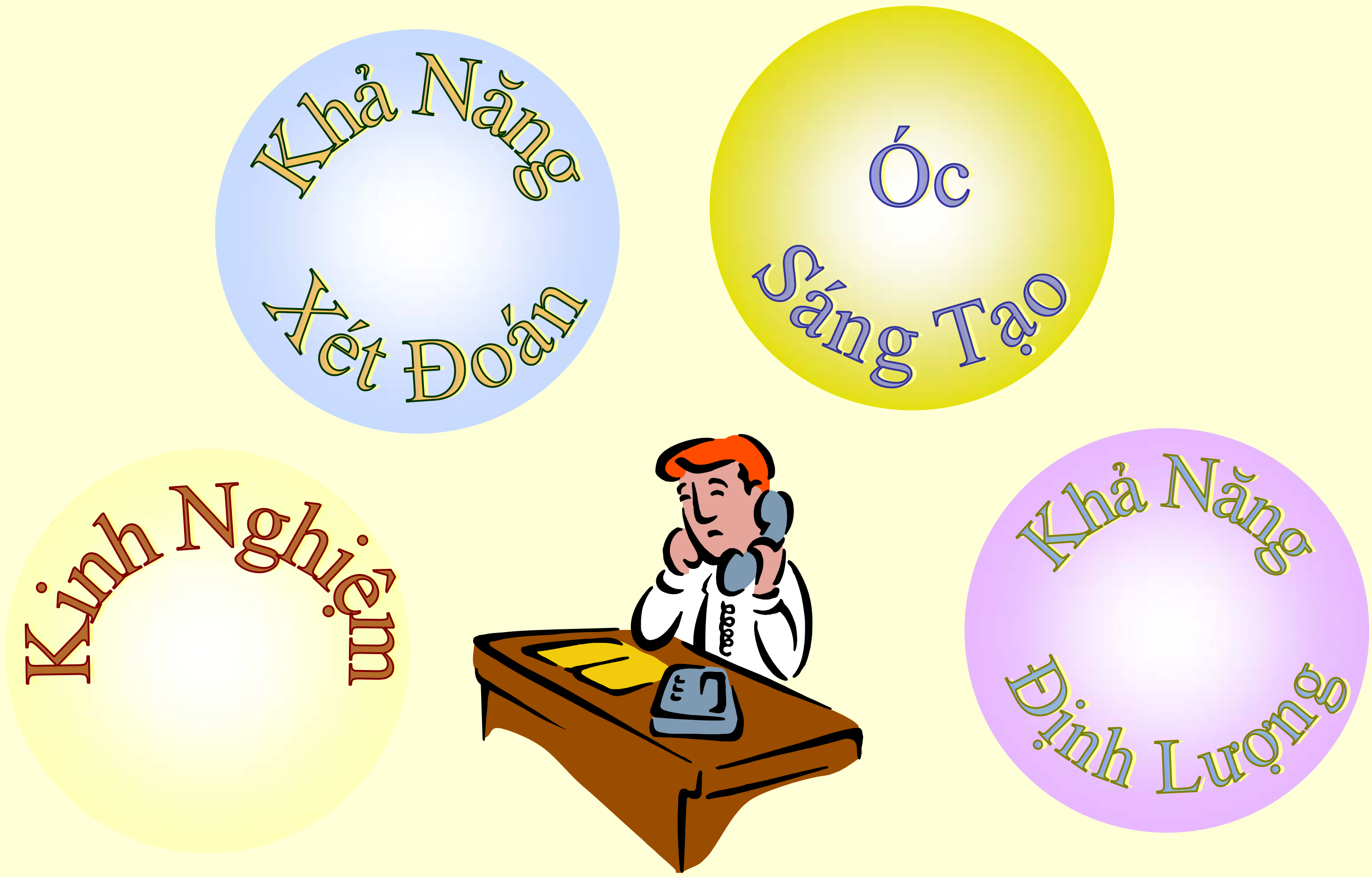
Nội dung

Chương 1

GIỚI THIỆU CHUNG

1.1. Phân tích định lượng và ra quyết định

Ra quyết định



1.1. Phân tích định lượng và ra quyết định

Tiến trình ra quyết định có thể được mô tả là một qui trình gồm 6 bước.

DECIDE

(1) Define the Problem (xác định vấn đề)

(2) Enumerate the decision factors (Liệt kê các yếu tố ảnh hưởng đến quyết định)

(3) Collect relevant information (Thu thập thông tin có liên quan)

(4) Identify the Solution (Quyết định giải pháp: gồm 3 bước nhỏ là đưa ra nhiều phương án khác nhau để lựa chọn, so sánh/đánh giá các phương án và lựa chọn phương án tốt nhất)

(5) Develop and Implement the solution (Tổ chức thực hiện quyết định)

(6) Evaluate the results (Đánh giá kết quả thực hiện quyết định)

1.1. Phân tích định lượng và ra quyết định

Quan điểm phân tích định lượng trong quản trị

- *Lý thuyết định lượng trong quản trị được xây dựng dựa trên nhận thức cơ bản rằng: "Quản trị là quyết định – (Management is decision making) và muốn việc quản trị có hiệu quả thì các quyết định phải đúng đắn"*
- Ra quyết định là nhiệm vụ quan trọng của nhà quản trị, kinh nghiệm, khả năng xét đoán, óc sáng tạo chưa thể đảm bảo có được những quyết định phù hợp và tối ưu nếu thiếu khả năng định lượng.
- Trong khi ra quyết định, nhà quản trị có thể sử dụng nhiều công cụ định lượng khác nhau với sự trợ giúp của máy tính.

1.1. Phân tích định lượng và ra quyết định

Quan điểm phân tích định lượng trong quản trị

- Chúng ta có thể mô tả qua sơ đồ sau:

CÁC CÔNG CỤ VÀ LÝ THUYẾT KINH TẾ
Lý thuyết về nhu cầu
Lý thuyết về doanh nghiệp
Lý thuyết sản xuất
Cơ cấu thị trường
Kinh tế học vĩ mô

CÁC CÔNG CỤ VÀ KHOA HỌC QUYẾT ĐỊNH
Các phương pháp thống kê
Dự báo và ước lượng
Tối ưu hóa
Các công cụ ra quyết định khoa học khác

KINH TẾ QUẢN LÝ
Sử dụng các công cụ và lý thuyết kinh tế cùng phương pháp luận khoa học trong việc ra quyết định để giải quyết các vấn đề kinh doanh và phân bổ nguồn lực tối ưu cho tổ chức

1.2. Nghiên cứu định lượng và định tính

Nghiên cứu định lượng và định tính

- Nghiên cứu định tính (NCĐT) là những nghiên cứu thu được các kết quả không sử dụng những công cụ đo lường, tính toán. Nói một cách cụ thể hơn NCĐT là những nghiên cứu tìm biết những đặc điểm, tính chất của đối tượng nghiên cứu (ĐTNC) cũng như những yếu tố ảnh hưởng đến suy nghĩ, hành vi của ĐTNC trong những hoàn cảnh cụ thể.
- Nghiên cứu định lượng (NCĐL) là những nghiên cứu thu được các kết quả bằng việc sử dụng những công cụ đo lường, tính toán với những con số cụ thể.
- Trong khi nghiên cứu định lượng (NCĐL) đi tìm trả lời cho câu hỏi bao nhiêu, mức nào (how many, how much) thì NCĐT đi tìm trả lời cho câu hỏi cái gì (what), như thế nào (how), tại sao (why). Ở một góc độ nào đó chính mục tiêu nghiên cứu là cơ sở để phân biệt nghiên cứu định lượng và định tính. Vì thế việc phát triển mục tiêu của một cuộc nghiên cứu là một bước hết sức quan trọng.

1.2. Nghiên cứu định lượng và định tính

Sự khác nhau cơ bản giữa NCĐL & NCĐT

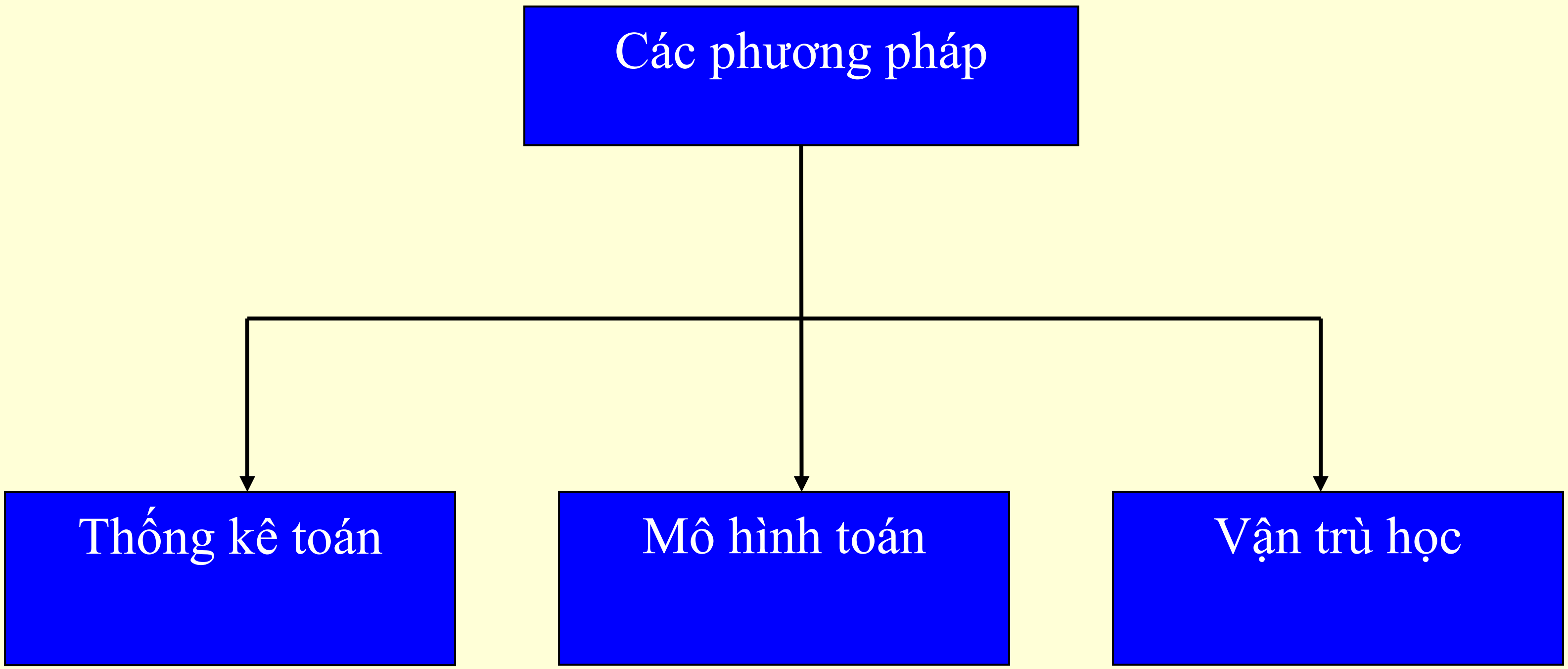
NCĐT	NCĐL
Dùng để mô tả, khám phá, thăm dò	Dùng để khẳng định, suy rộng và dự báo
Chỉ tiêu, đối tượng NC, mức độ nghiên cứu có thể chưa rõ ràng	Chỉ tiêu, đối tượng NC, mức độ nghiên cứu đã rõ ràng
Linh động trong hướng nghiên cứu, khám phá các hướng nghiên cứu chưa biết	Yêu cầu phải đo lường
Người nghiên cứu là công cụ thu thập thông tin	Người nghiên cứu sử dụng các công cụ như bản câu hỏi để thu thập thông tin
Người nghiên cứu biết sơ bộ những điều mà họ muốn nghiên cứu	Người nghiên cứu biết rõ ràng những điều mà họ muốn nghiên cứu
Chủ quan: Ý kiến của cá nhân là quan trọng, vd: quan sát, phỏng vấn sâu	Khách quan: đo lường và phân tích qua điều tra
Quy nạp giả thuyết	Kiểm tra giả thuyết
Khó khái quát hóa	Khái quát hóa
Từ ngữ, hình ảnh	Con số, thống kê

1.3. Mục tiêu của nghiên cứu định lượng

- Khẳng định, suy rộng và dự báo,
- Để nhận dạng vấn đề,
- Kiểm định một lý thuyết hay một giả thiết,
- Đo lường các con số, và phân tích bằng các kỹ thuật thống kê,
- Lập kế hoạch sản xuất
- Để tính toán lựa chọn phương án tối ưu (Quyết định đầu tư, lựa chọn các phương án qui hoạch…)

1.4. Phương pháp và các bước tiến hành

Các phương pháp toán ứng dụng trong phân tích định lượng



1.4. Phương pháp và các bước tiến hành

Các phương pháp toán ứng dụng trong phân tích định lượng

- ***Thống kê kế toán:*** Là một bộ phận của toán học ứng dụng dành cho các phương pháp xử lý và phân tích số liệu thống kê, mà các ứng dụng chủ yếu của nó trong quản lý là các phương pháp xử lý kiểm tra và dự đoán (dự đoán, điều tra chọn mẫu,…)
- ***Mô hình toán:*** Là sự phản ánh những thuộc tính cơ bản nhất định của các đối tượng nghiên cứu kinh tế, là công cụ trọng cho việc trừu tượng hoá một cách khoa học các quá trình và hiện tượng kinh tế.

  Khoa học kinh tế từ lâu đã biết sử dụng các mô hình kinh tế lượng như mô hình hàm sản suất Cobb – Douglas, mô hình cung cầu, giá cả v.v...

1.4. Phương pháp và các bước tiến hành

Các phương pháp toán ứng dụng trong phân tích định lượng

- ***Vận trù học:*** Là khoa học có mục đích nghiên cứu các phương pháp phân tích nhằm chuẩn bị căn cứ chính xác cho các quyết định, đối tượng của nó là hệ thống, tức là tập hợp các phần tử và hệ thống còn có tác động qua lại với nhau nhằm đạt tới một mục tiêu nhất định. Vận trù học bao gồm nhiều nhánh khoa học ứng dụng gộp lại: (1) Lý thuyết tối ưu (bao gồm: quy hoạch tuyến tính, quy hoạch động, quy hoạch ngẫu nhiên, quy hoạch nguyên, quy hoạch 0 – 1, quy hoạch đa mục tiêu, lý thuyết trò chơi...); (2) Lý thuyết đồ thị và sơ đồ mạng lưới; (3) Lý thuyết dự trữ bảo quản; (4) Lý thuyết tìm kiếm;...

1.4. Phương pháp và các bước tiến hành

Các phương pháp toán ứng dụng trong phân tích định lượng

Các phương pháp và mô hình cơ bản:

- Thống kê mô tả
- Phương pháp Phân tích hồi qui,
- Các phương pháp Dự báo,
- Mô hình toán (qui hoạch tuyến tính, qui hoạch nguyên, qui hoạch phi tuyến),
- Mô hình mạng,
- Phân tích Markov,…

1.4. Phương pháp và các bước tiến hành

Các bước tiến hành phân tích định lượng

Xác định vấn đề

↓

Thu thập dữ liệu

↓

Xây dựng mô hình

↓

Tính toán

↓

Phân tích kết quả

↓

Áp dụng kết quả

1.5. Các phần mềm ứng dụng

- EXCEL
- SPSS
- EVIEWS
- LINDO, LINGO.

Chương 2

PHÂN PHỐI XÁC SUẤT VÀ THỐNG KÊ TOÁN

Nội dung

2.1. Biến ngẫu nhiên

2.2. Đo lường sự định tâm

2.3. Đo lường sự biến thiên và tương quan

2.4. Phân phối xác suất

2.5. Ước lượng thống kê

2.6. Kiểm định giả thiết thống kê

2.1. Biến ngẫu nhiên

Định nghĩa

- **"Một biến ngẫu nhiên là một qui tắc hay một hàm số để gán các giá trị bằng số cho những kết quả của một trắc nghiệm ngẫu nhiên."**
- Các biến ngẫu nhiên thường được ký hiệu bằng các chữ lớn X, Y, Z,… còn các giá trị của chúng được ký hiệu bằng các chữ nhỏ x, y, z...

2.1. Biến ngẫu nhiên

Phân loại

- **Biến ngẫu nhiên rời rạc** (Discrete Random Variable)
  - Nếu giá trị của biến ngẫu nhiên X có thể lập thành dãy rời rạc các số $x_1, x_2, \ldots, x_n$ (dãy hữu hạn hay vô hạn) thì X được gọi là biến ngẫu nhiên rời rạc.
  - Trắc nghiệm: thảy hai xúc xắc và tính tổng. Trắc nghiệm ngẫu nhiên bao gồm việc thảy xúc xắc này. Nhà nghiên cứu tính xem xuất hiện bao nhiêu chấm trên mặt từng xúc xắc và tính chúng. Dựa trên trắc nghiệm này chúng ta có thể xác định nhiều biến ngẫu nhiên.
  - Gọi X_1 là số các chấm thể hiện trên xúc xắc thứ nhất. Những kết quả có thể có của biến ngẫu nhiên X_1 này là { 1, 2, 3, 4, 5, 6 }.
  - Gọi X_2 là số các chấm thể hiện trên xúc xắc thứ hai. Những kết quả có thể có của biến ngẫu nhiên X_2 này là { 1, 2, 3, 4, 5, 6 }.
  - Đặt $X = X_1 + X_2$. Những kết quả có thể có của biến ngẫu nhiên này là {2, ..., 12}.

2.1. Biến ngẫu nhiên

Phân loại

- **Biến ngẫu nhiên liên tục** (Continuous Random Variable)
  - Nếu giá trị của biến ngẫu nhiên X có thể lấp đầy toàn bộ khoảng hữu hạn hay vô hạn (a,b) của trục số 0x thì biến ngẫu nhiên X được gọi là biến ngẫu nhiên liên tục.
  - Nếu chúng ta nghĩ về tiếp cận tần suất tương đối tới xác suất, và chúng ta tưởng tượng việc lựa chọn một quan sát ngẫu nhiên, dường như rõ ràng là xác suất của việc thu được *chính xác* một giá trị nhất định phải là zero. Mặt khác, nếu chúng ta đặt vấn đề dưới dạng khoảng, thì việc xác định xác suất này là đơn giản.
  - Hãy tưởng tượng rằng đang mưa và rằng Anh/Chị đặt một thước đo trên mặt đất. Xác suất để hạt mưa sau sẽ rơi vào giữa 0 và 10 cm là gì? Xác suất để hạt mưa sau sẽ rơi vào giữa 10 và 20 cm là gì?
  - Chúng ta có thể chia thước đo này thành 10 bước với khoảng cách là 10 cm mỗi bước. Xác suất để một hạt mưa rơi vào bất cứ khoảng cụ thể nào sẽ bằng 1/k, trong đó k là số các khoảng trong thước. Trong trường hợp này, việc tính xác suất để một hạt mưa rơi vào một khoảng có bất cứ độ dài cụ thể nào thì thật là đơn giản.

2.2. Đo lường sự định tâm

Kỳ vọng toán học của 1 biến ngẫu nhiên (số trung bình)

- Định nghĩa: Cho X là 1 biến ngẫu nhiên, giá trị trung bình hay kỳ vọng toán học (gọi tắt là kỳ vọng) của X được ký hiệu là EX và được tính theo công thức:

$$EX = \mu = \sum_x x.P(x) \qquad \text{Nếu x rời rạc}$$

$$= \mu = \int_{-\infty}^{+\infty} xf(x)dx \qquad \text{Nếu x liên tục}$$

- ***Chú ý***: Nếu mẫu ngẫu nhiên cho dưới dạng tần suất:

X	X_1	X_2	X_3	...	X_k
n_i	n_1	n_2	n_3	...	n_k

thì trung bình mẫu được tính:

$$\overline{X} = \frac{n_1X_1 + n_2X_2 + n_3X_3 + \ldots + n_kX_k}{n_1 + n_2 + n_3 + \ldots + n_k} = \frac{\sum_{i=1}^{k} n_iX_i}{\sum_{i=1}^{k} n_i}$$

2.2. Đo lường sự định tâm

Kỳ vọng toán học của 1 biến ngẫu nhiên (số trung bình)

- **Ví dụ 1**: Cho mẫu quan sát (Xi) với i = 1, 2, ..., 10 của ĐLNN X là:

X_i	1	2	3	4	5	6	7	8	9	10
n_i	3	7	5	6	5	8	4	2	6	4

- Khi đó: Trung bình mẫu của ĐLNN X là

$$\overline{X} = \frac{3.1 + 7.2 + 5.3 + 6.4 + 5.5 + 8.6 + 4.7 + 2.8 + 6.9 + 4.10}{50} = \frac{273}{50} = 5,46$$

Kỳ vọng toán học của 1 biến ngẫu nhiên (số trung bình)

- **Ví dụ 2:** Giả sử X là số xe máy đến 1 cửa hàng rửa xe vào chiều thứ 7 hàng tuần có bảng phân bố xác xuất:

X	4	5	6	7	8	9
P(x)	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

Tìm kỳ vọng EX của biến ngẫu nhiên X (số xe máy trung bình tới trạm rửa xe vào chiều thứ 7).

Giải:

$$EX = 4.\frac{1}{12} + 5.\frac{1}{12} + 6.\frac{1}{4} + 7\frac{1}{4} + 8.\frac{1}{6} + 9.\frac{1}{6} = \frac{89}{12}$$

2.2. Đo lường sự định tâm

Kỳ vọng toán học của 1 biến ngẫu nhiên (số trung bình)

- **<u>Các qui tắc</u>**:

1. $E(X+Y) = E(X) + E(Y)$

Suy rộng:

$E(W+X+Y+Z) = E(W) + E(X) + E(Y) + E(Z)$

2. $E(bX) = bE(X)$

Ví dụ: $E(3X) = 3E(X)$

3. $E(b) = b$

Số trung vị, số yếu vị

- **Số trung vị (Median)**

  Số trung vị của khối Dữ liệu là số mà phân nửa giá trị quan sát được của khối Dữ liệu nhỏ hơn nó và phân nữa giá trị quan sát lớn hơn nó.

  Gọi n là số giá trị quan sát được (đối với biến ngẫu nhiên rời rạc)

  - Nếu n là số lẻ thì số trung vị là số có thứ tự (n+1)/2. Nó chính là số có vị trí ở giữa khối Dữ liệu.
  - Nếu n là số chẵn thì số trung vị là trung bình cộng của hai số có thứ tự n/2 và n/2+1

- **Số yếu vị (Mode)**

  Số yếu vị của khối Dữ liệu là số có tần số lớn nhất

Ví dụ

- Cho khối dữ kiện: 0 1 0 2 5 2 5 2 3 3 5 6 4

  Tìm số trung bình, số trung vị và số yếu vị của khối Dữ liệu.
- **Giải**:

  Ta có bảng phân phối tần số :

X	0	1	2	3	4	5	6
Tần số f_i	2	1	3	2	1	3	1

- Số trung bình (Mean) $$\overline{X} = \frac{\sum_{i=1}^{7} f_i x_i}{\sum_{i=1}^{7} f_i} = \frac{2x0 + 1x1 + 3x2 + 2x3 + 1x4 + 3x5 + 1x6}{13} = 2{,}923$$
- Số trung vị (Median): Cỡ mẫu n = 13 lẻ => (n+1)/2 = 7

  0 0 1 2 2 2 3 3 4 5 5 5 6

  □ Số trung vị là số có thứ tự 7, nghĩa là số trung vị là 3
- Số yếu vị là 2 và 5 có tần số lớn nhất là 3

2.3. Đo lường sự biến thiên và tương quan

Phương sai và Covariance (hiệp phương sai)

Phương sai:

- Định nghĩa: Nếu X có kỳ vọng EX = μ thì phương sai của X ký hiệu là σ^2 hay DX được tính theo công thức:

$$\sigma^2 = E(X-\mu)^2 = \sum_x (x-\mu)^2 P(x) \quad \text{nếu X rời rạc}$$

$$\sigma^2 = E(X-\mu)^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)\,\mathrm{dx} \quad \text{nếu X liên tục}$$

- Chú ý: Căn bậc hai của phương sai, σ gọi là độ lệch chuẩn của X
- Định lý: Phương sai của biến ngẫu nhiên X còn được tính theo công thức: $\sigma^2 = E(X)^2 - \mu^2$
- **Ý nghĩa:** Phương sai đo sự phân tán của các giá trị của X quanh kỳ vọng của nó.
- Phương sai mẫu được tính như sau: $S_X^2 = \dfrac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}$

2.3. Đo lường sự biến thiên và tương quan

Phương sai và Covariance (hiệp phương sai)

- VD: Cho X là số xe ô tô được sử dụng vào 1 mục đích phục vụ đào tạo của 1 trường đại học. Giả sử X có phân bố:

X	1	2	3
P(x)	0,3	0,4	0,3

Tìm EX và DX

Giải: $\mu = E(X) = 1.\ (0{,}3) + 2.\ (0{,}4) + 3.\ (0{,}3) = 2$

$$DX = \sum_{i=1}^{3} (Xi - EX)^2 Pi = (1-2)^2(0{,}3) + (2-2)^2(0{,}4) + (3-2)^2(0{,}3) = 0{,}6$$

- Chú ý: Có thể tính DX theo công thức: $DX = EX^2 - (EX)^2$

Phương sai và Covariance (hiệp phương sai)

Hiệp phương sai:

- Định nghĩa: Cho (X, Y) là 2 biến ngẫu nhiên, Covariance của X và Y được ký hiệu là σ_{XY} và tính theo công thức:

$$\sigma_{XY} = E[(X - EX)(Y - EY)]$$

- Nếu EX = μ_X, EY = μ_Y, Covariance của X và Y còn có thể tính theo công thức: $\sigma_{XY} = E(XY) - \mu_X . \mu_Y$
- Đối với biến ngẫu nhiên rời rạc:

$$\operatorname{cov}(X,Y) = \sum_y \sum_x (X - \mu_x)(Y - \mu_y) f(x,y) = \sum_y \sum_x XYf(x,y) - \mu_x \mu_y$$

- Đối với biến ngẫu nhiên liên tục:

$$\operatorname{cov}(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (X - \mu_x)(Y - \mu_y) f(x,y)dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} XYf(x,y)dxdy - \mu_x \mu_y$$

Phương sai và Covariance (hiệp phương sai)

Hệ số tương quan:

- Để khảo sát sự phụ thuộc hay mức độ độc lập của 2 biến ngẫu nhiên X, Y và khắc phục nhược điểm của hiệp phương sai là phụ thuộc vào đơn vị đo lường, người ta sử dụng hệ số tương quan được định nghĩa như sau:

$$\rho_{xy} = \frac{\operatorname{cov}(X,Y)}{\sqrt{\operatorname{var}(X)\operatorname{var}(Y)}} = \frac{\operatorname{cov}(X,Y)}{\sigma_x \sigma_y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \qquad 0 \le |\rho_{XY}| \le 1$$

- Hệ số tương quan đo lường mối quan hệ tuyến tính giữa hai biến. ρ sẽ nhận giá trị nằm giữa -1 và 1. Nếu ρ = -1 thì mối quan hệ là nghịch biến hoàn hảo, nếu ρ = 1 thì mối quan hệ là đồng biến hoàn hảo.

Phương sai và Covariance (hiệp phương sai)

Một số qui tắc của Phương sai:

- Qui tắc 1:
  Nếu $Y = V + W$, $\mathrm{Var}(Y) = \mathrm{Var}(V) + \mathrm{Var}(W) + 2\mathrm{Cov}(V, W)$
- Qui tắc 2:
  Nếu $Y = bZ$, trong đó b là hằng số, $\mathrm{Var}(Y) = b^2\mathrm{Var}(Z)$
- Qui tắc 3:
  Nếu $Y = b$, trong đó b là hằng số, $\mathrm{Var}(Y) = 0$
- Qui tắc 4:
  Nếu $Y = V + b$, trong đó b là hằng số, $\mathrm{Var}(Y) = \mathrm{Var}(V)$

2.3. Đo lường sự biến thiên và tương quan

Phương sai và Covariance (hiệp phương sai)

Một số qui tắc của Covariance:

- Qui tắc 1:
  Nếu $Y = V + W$, $\text{Cov}(X, Y) = \text{Cov}(X, V) + \text{Cov}(X, W)$
- Qui tắc 2:
  Nếu $Y = bZ$, trong đó b là hằng số, $\text{Cov}(X, Y) = \text{Cov}(X, bZ) = b\text{Cov}(X, Z)$
- Qui tắc 3:
  Nếu $Y = b$, trong đó b là hằng số, $\text{Cov}(X, Y) = \text{Cov}(X, b) = 0$

2.4. Phân phối xác suất

Khái niệm

- Mỗi biến ngẫu nhiên tạo ra một phân phối xác suất, phân phối này chứa hầu hết các thông tin quan trọng về biến ngẫu nhiên đó. Nếu X là một biến ngẫu nhiên, phân phối xác suất tương ứng gán cho đoạn $[a, b]$ một xác suất $P[a \leq X \leq b]$, nghĩa là, xác suất mà biến X sẽ lấy giá trị trong đoạn $[a, b]$.
- Phân phối xác suất của biến X có thể được mô tả bởi hàm phân phối tích lũy (*cumulative distribution function*) $F(x)$ được định nghĩa như sau:

$$F(x) = P[\, X \leq x \,]$$

2.4. Phân phối xác suất

Khái niệm

- Một phân phối được gọi là *rời rạc* nếu hàm phân phối tích lũy của nó bao gồm một dãy các bước nhảy hữu hạn, nghĩa là nó sinh ra từ một biến ngẫu nhiên rời rạc *X*: một biến chỉ có thể nhận giá trị trong một tập hợp hữu hạn hoặc đếm được nhất định. Một phân phối được gọi là *liên tục* nếu hàm phân phối tích lũy của nó là hàm liên tục, khi đó nó sinh ra từ một biến ngẫu nhiên *X* mà P[$X = x$] = 0 với mọi x thuộc **R**. Phân phối liên tục còn có thể được biểu diễn bằng hàm mật độ xác suất như sau:

$$\boxed{P(a < x < b) = \int_a^b f(x)dx}$$

2.4. Phân phối xác suất

Khái niệm- ví dụ phân phối xác suất rời rạc

Thí dụ

Trong thí nghiệm thảy 1 con xúc xắc, ta có:

P(X=1) = P(X=2) = … = P(X=6) = 1/6

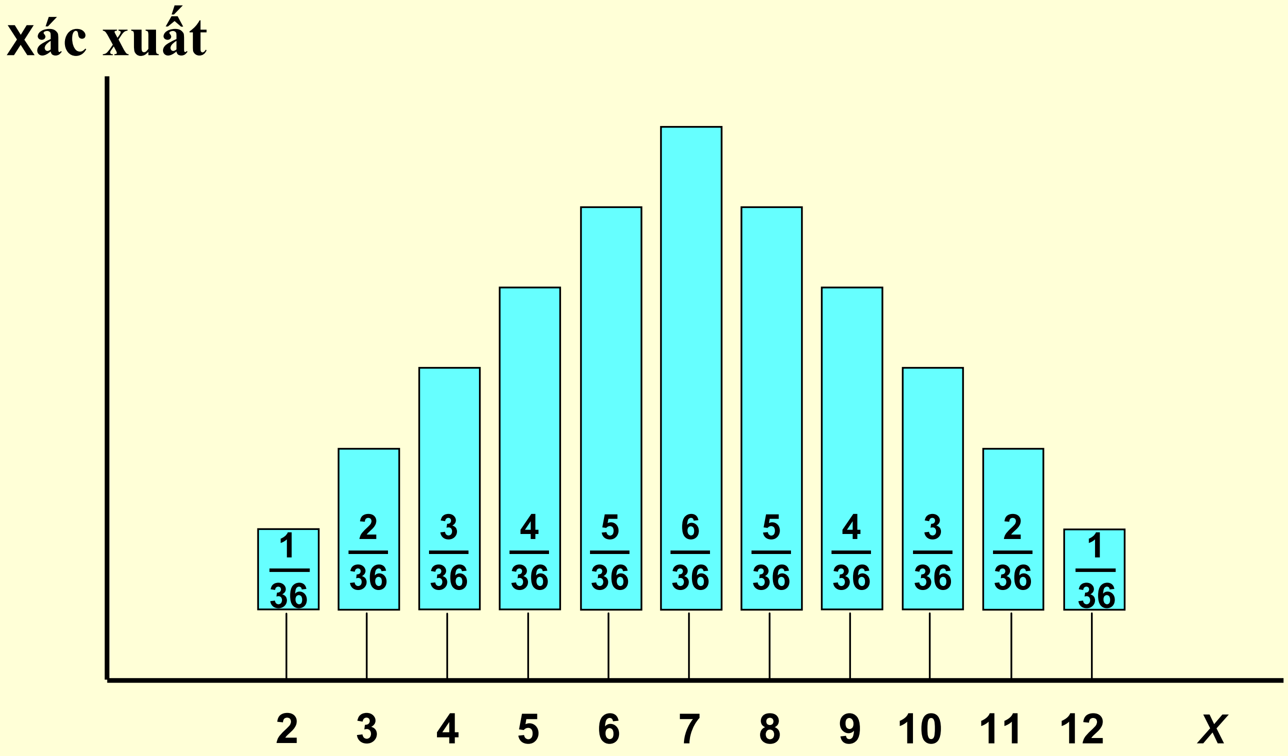
=> Hàm xác suất là: PX(x) = P(X=x) = 1/6 với x =1, 2, 3, 4, 5, 6

Đỏ / Xanh	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

X	*f*	*p*
2	1	1/36
3	2	2/36
4	3	3/36
5	4	4/36
6	5	5/36
7	6	6/36
8	5	5/36
9	4	4/36
10	3	3/36
11	2	2/36
12	1	1/36

2.4. Phân phối xác suất

Khái niệm- ví dụ phân phối xác suất rời rạc

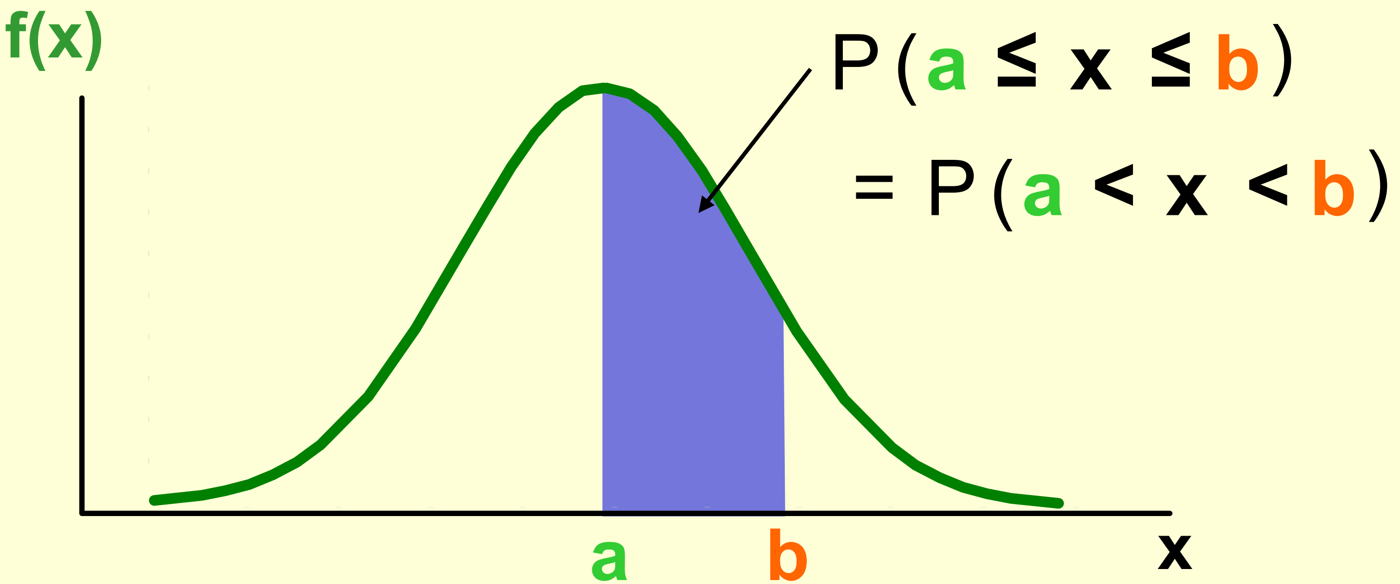


Phân phối được thể hiện bằng đồ thị. Trong ví dụ này nó đối xứng, xác suất xảy ra cao nhất đối với X bằng 7.

2.4. Phân phối xác suất

Khái niệm- ví dụ phân phối xác suất liên tục

Phân phối xác suất của X trong khoảng a, b



2.4. Phân phối xác suất

Một số phân phối xác suất thường dùng

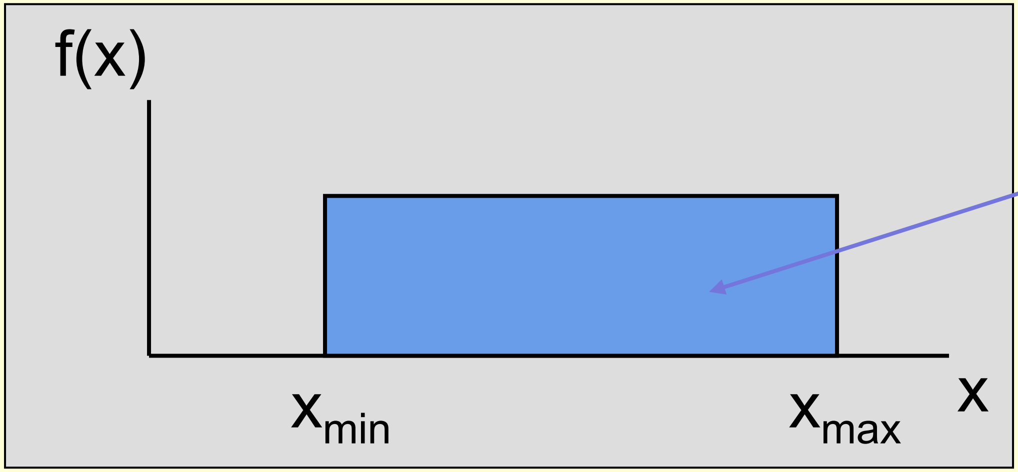
1. Uniform Distribution/ Phân phối đều liên tục
2. Normal Distribution/ Phân phối chuẩn
3. z-Distribution/ Phân phối chuẩn hoá
4. t-Distribution/ Phân phối T
5. F-Distribution/ Phân phối F
6. Chi-Square Distribution/ Phân phối chi bình phương

2.4. Phân phối xác suất

Một số phân phối xác suất thường dùng

1. Uniform Distribution/ Phân phối đều liên tục

- **Phân phối đều liên tục** là một phân phối mà xác suất xảy ra như nhau cho mọi kết cục của biến ngẫu nhiên liên tục. Phân phối đều liên tục đôi khi còn được gọi là phân phối **hình chữ nhật** và khi biểu diễn bằng hình vẽ sẽ có dạng hình chữ nhật.



Tổng xác suất trong toàn bộ miền hình chữ nhật bằng 1.0

Một số phân phối xác suất thường dùng

1. Uniform Distribution/ Phân phối đều liên tục

- Hàm mật độ xác suất của một phân phối đều liên tục có dạng:

$$f(x) = \begin{cases} \dfrac{1}{b-a} & ; a \le x \le b \\ 0 & ; x < a \text{ hay } x > b \end{cases}$$

Trong đó: x là biến ngẫu nhiên liên tục, a là giá trị cực tiểu, b là giá trị cực đại.

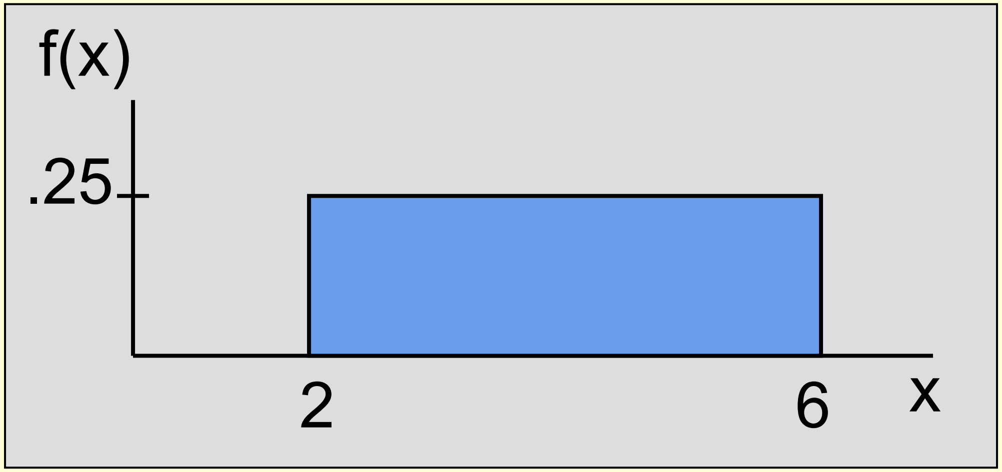
Giá trị kỳ vọng là: $\mu = \dfrac{a+b}{2}$ *Phương sai là:* $\sigma^2 = \dfrac{(b-a)^2}{12}$

Một số phân phối xác suất thường dùng

1. Uniform Distribution/ Phân phối đều liên tục

- Ví dụ: Phân phối xác suất trong khoảng $2 \le x \le 6$:

$$f(x) = \frac{1}{6-2} = .25 \quad \text{với} \quad 2 \le x \le 6$$



$$\mu = \frac{a+b}{2} = \frac{2+6}{2} = 4$$

$$\sigma^2 = \frac{(b-a)^2}{12} = \frac{(6-2)^2}{12} = 1.333$$

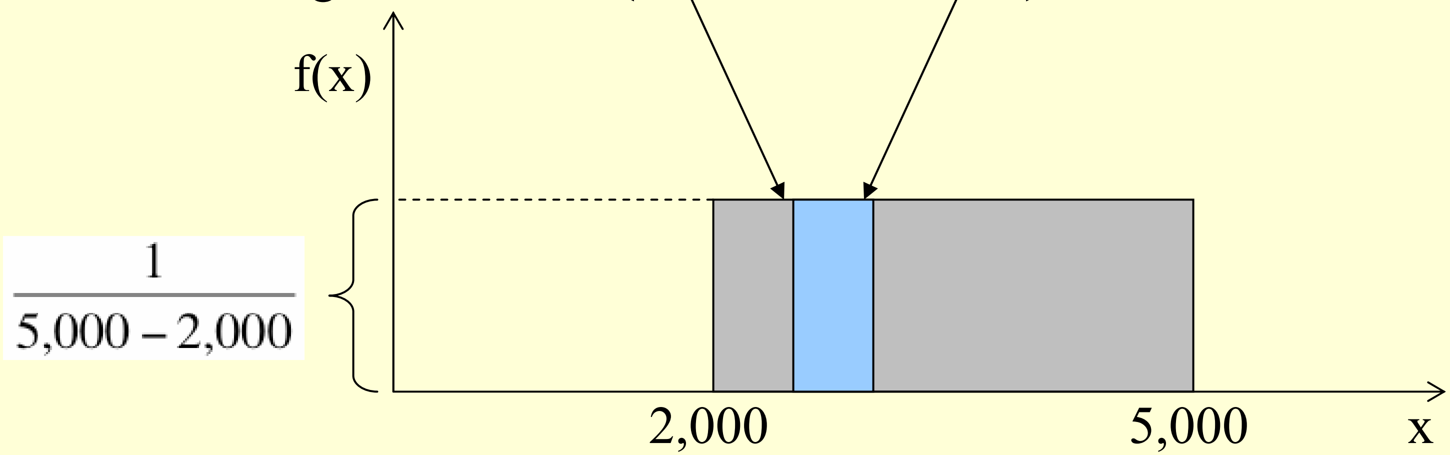
2.4. Phân phối xác suất

Một số phân phối xác suất thường dùng

1. Uniform Distribution/ Phân phối đều liên tục

- Ví dụ: Lượng xăng bán hàng ngày ở một cửa hàng tối thiểu là 2,000 lít và tối đa là 5,000 lít, Tìm xác suất bán trong ngày nằm trong khoảng 2,500 đến 3,000 lít.

Có nghĩa là: Tìm **P(2,500 ≤ X ≤ 3,000)** ?



- Giải:

$$P(2{,}500 \le X \le 3{,}000) = (3{,}000 - 2{,}500) * \frac{1}{3{,}000} = 0.1667$$

=> Xác suất bán một ngày trong khoảng 2,500 đến 3,000 lít là 17%

2.4. Phân phối xác suất

Một số phân phối xác suất thường dùng

2. Normal Distribution/ Phân phối chuẩn

- **Phân phối chuẩn**, còn gọi là phân phối **Gauss**, là một phân phối xác suất cực kì quan trọng trong nhiều lĩnh vực. Nó là họ phân phối có dạng tổng quát giống nhau, chỉ khác tham số *vị trí* (giá trị trung bình μ) và *tỉ lệ* (phương sai σ^2).
- Hàm phân phối được xác định như sau:

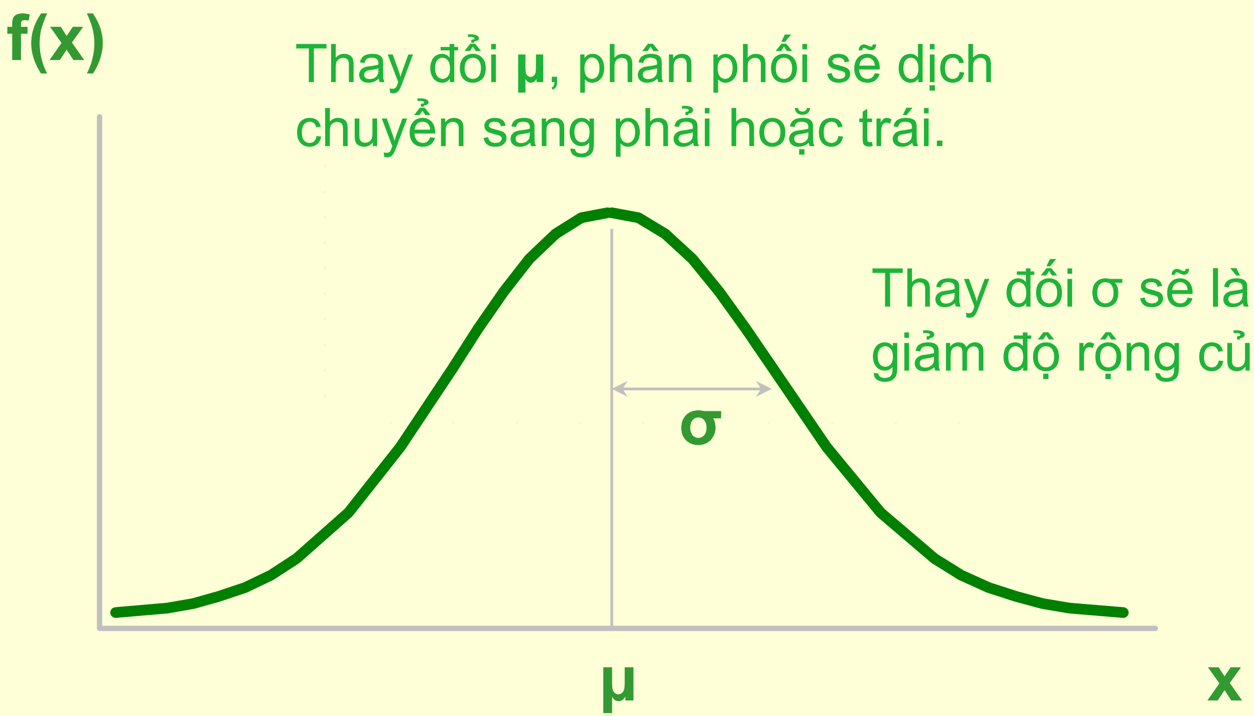
$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(x-\mu)^2/2\sigma^2}$$

Trong đó:

- $e = 2.71828$
- $\pi = 3.14159$
- μ = giá trị kỳ vọng
- σ = Độ lệch chuẩn
- x = Giá trị bất kỳ của biến, $-\infty < x < \infty$

Một số phân phối xác suất thường dùng

2. Normal Distribution/ Phân phối chuẩn

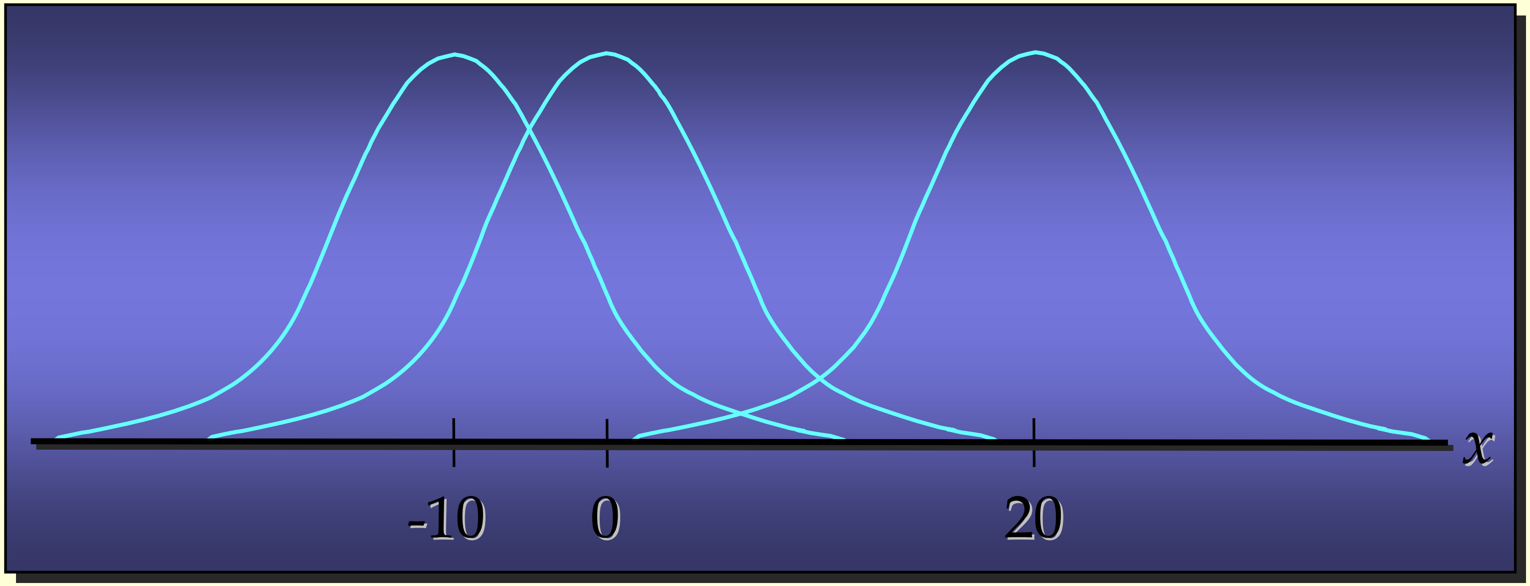


Ký hiệu phân phối chuẩn $\boxed{X \sim N(\mu, \sigma^2)}$

Một số phân phối xác suất thường dùng

2. Normal Distribution/ Phân phối chuẩn

Phân phối chuẩn có phương sai bằng nhau nhưng kỳ vọng khác nhau

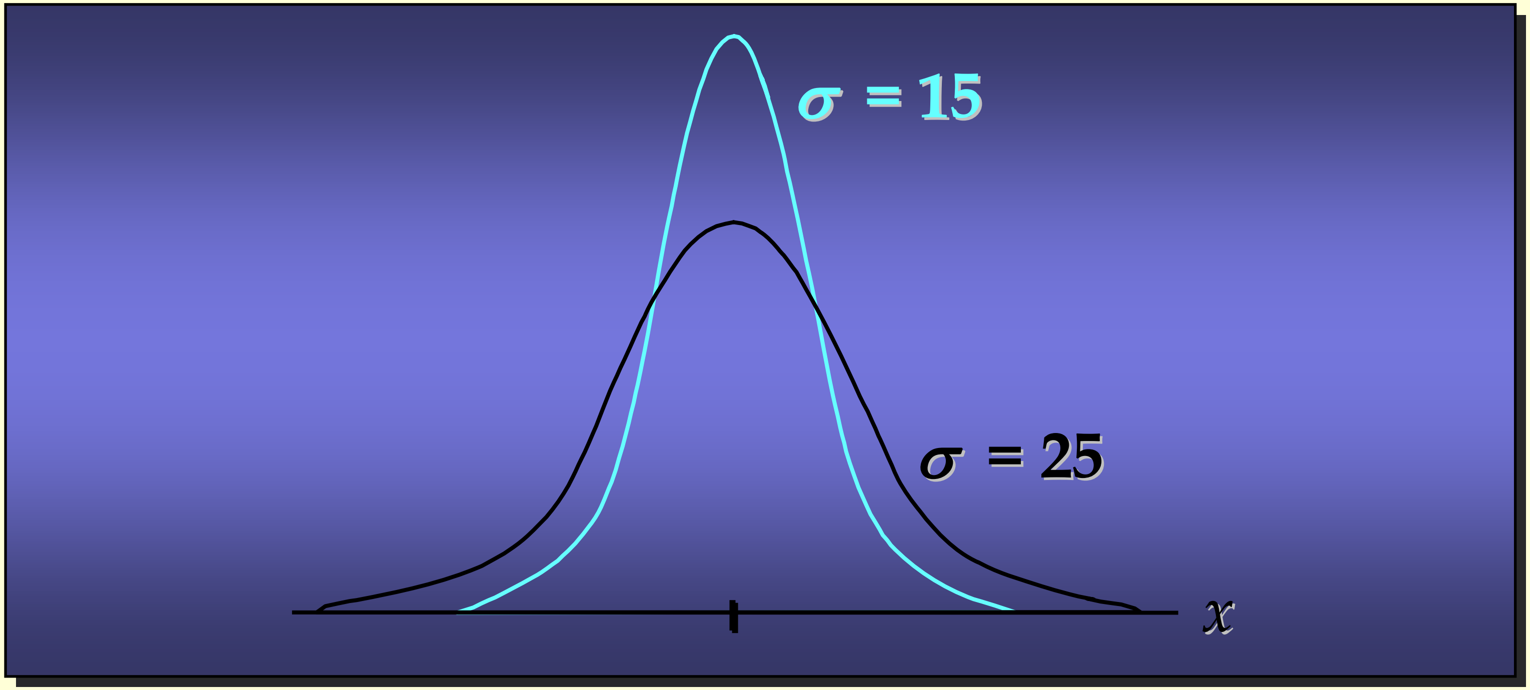


2.4. Phân phối xác suất

Một số phân phối xác suất thường dùng

2. Normal Distribution/ Phân phối chuẩn

Phân phối chuẩn có phương sai khác nhau nhưng kỳ vọng bằng nhau



Một số phân phối xác suất thường dùng

2. Normal Distribution/ Phân phối chuẩn

- **Tính chất của phân phối chuẩn**
  - Hàm mật độ xác suất của đối xứng quanh giá trị trung bình.
  - Xấp xỉ 68% diện tích dưới đường phân phối (pdf-probability density function) nằm trong khoảng $\mu\pm\sigma$, xấp xỉ 95% diện tích nằm dưới đường pdf nằm trong khoảng $\mu\pm2\sigma$, và xấp xỉ 99,7% diện tích nằm dưới đường pdf nằm trong khoảng $\mu\pm3\sigma$.
  - Định lý giới hạn trung tâm 1: Một kết hợp tuyến tính các biến có phân phối chuẩn,, trong một số điều kiện xác định cũng là một phân phối chuẩn. Ví dụ X_1 và X_2 là 2 biến ngẫu nhiên có phân phối chuẩn thì $Y = aX_1+bX_2$ với a và b là hằng số có phân phối $Y \sim N[(a\mu_1+b\mu_2),(a\sigma_1^2+b\sigma_2^2)]$.
  - Định lý giới hạn trung tâm 2: Dưới điều kiện xác định, giá trị trung bình mẫu của các một biến ngẫu nhiên sẽ gần như tuân theo phân phối chuẩn.

2.4. Phân phối xác suất

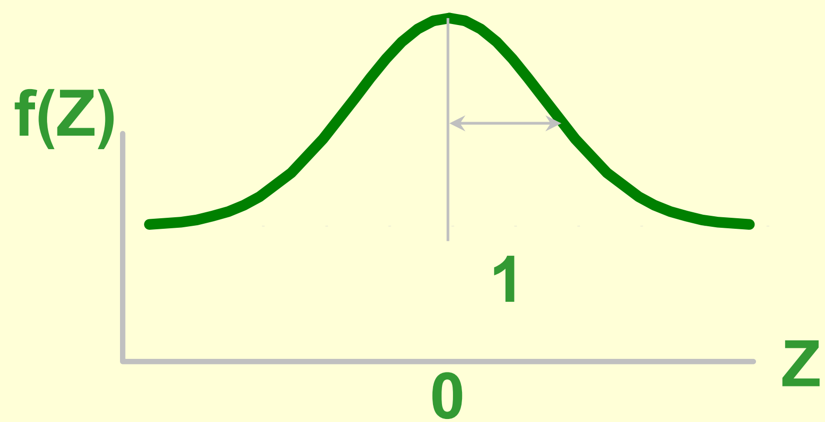
Một số phân phối xác suất thường dùng

3. z-Distribution/ Phân phối chuẩn hoá

- Phân phối chuẩn hóa (*standard normal distribution*) là phân phối chuẩn với giá trị trung bình bằng 0 và phương sai bằng 1. Nếu đặt Z = (X-μ)/σ thì ta có Z~N(0,1). Z gọi là biến chuẩn hoá và N(0,1) được gọi là phân phối chuẩn hoá

$$Z \sim N(0,1)$$

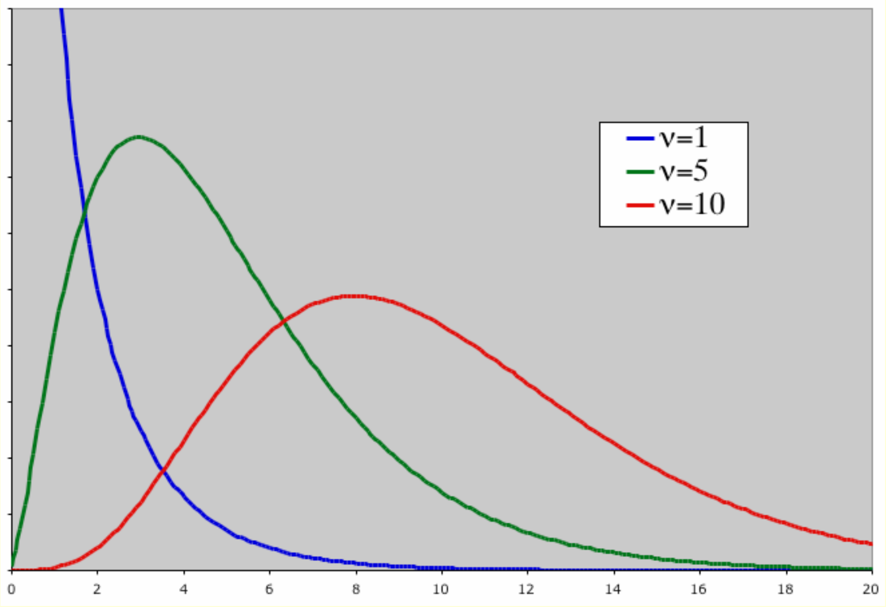
$$Z = \frac{X - \mu}{\sigma}$$



Một số phân phối xác suất thường dùng

4. Chi-Square Distribution/ Phân phối chi-bình phương

- Giả sử z_1, z_2, ... z_k là k biến ngẫu nhiên độc lập thống kê và có phân phối chuẩn hoá. Người ta nói rằng tổng bình phương của các biến ngẫu nhiên đó sẽ tuân theo phân phối Chi-bình phương với n là bậc tự do. Được ký hiệu là: (χ^2).



Một số phân phối xác suất thường dùng

4. Chi-Square Distribution/ Phân phối chi-bình phương

- Tính chất của phân phối Chi-bình phương.
  - Phân phối chi bình phương bắt đầu từ gốc tọa độ, lệch về phía bên trái và có đuôi dài vô tận về phía phải. Khi bậc tự do tăng dần thì phân phối χ^2 tiến gần đến phân phối chuẩn.
  - $\mu = k$ và $\sigma^2 = 2k$
  - $\chi^2_{k1} + \chi^2_{k2} = \chi^2_{k1+k2}$ hay tổng của hai biến có phân phối χ^2 cũng có phân phối χ^2 với số bậc tự do bằng tổng các bậc tự do.

Một số phân phối xác suất thường dùng

5. t-Distribution/ Phân phối T

- Nếu Z~N(0,1) và χ^2 có phân phối Chi-bình phương thì

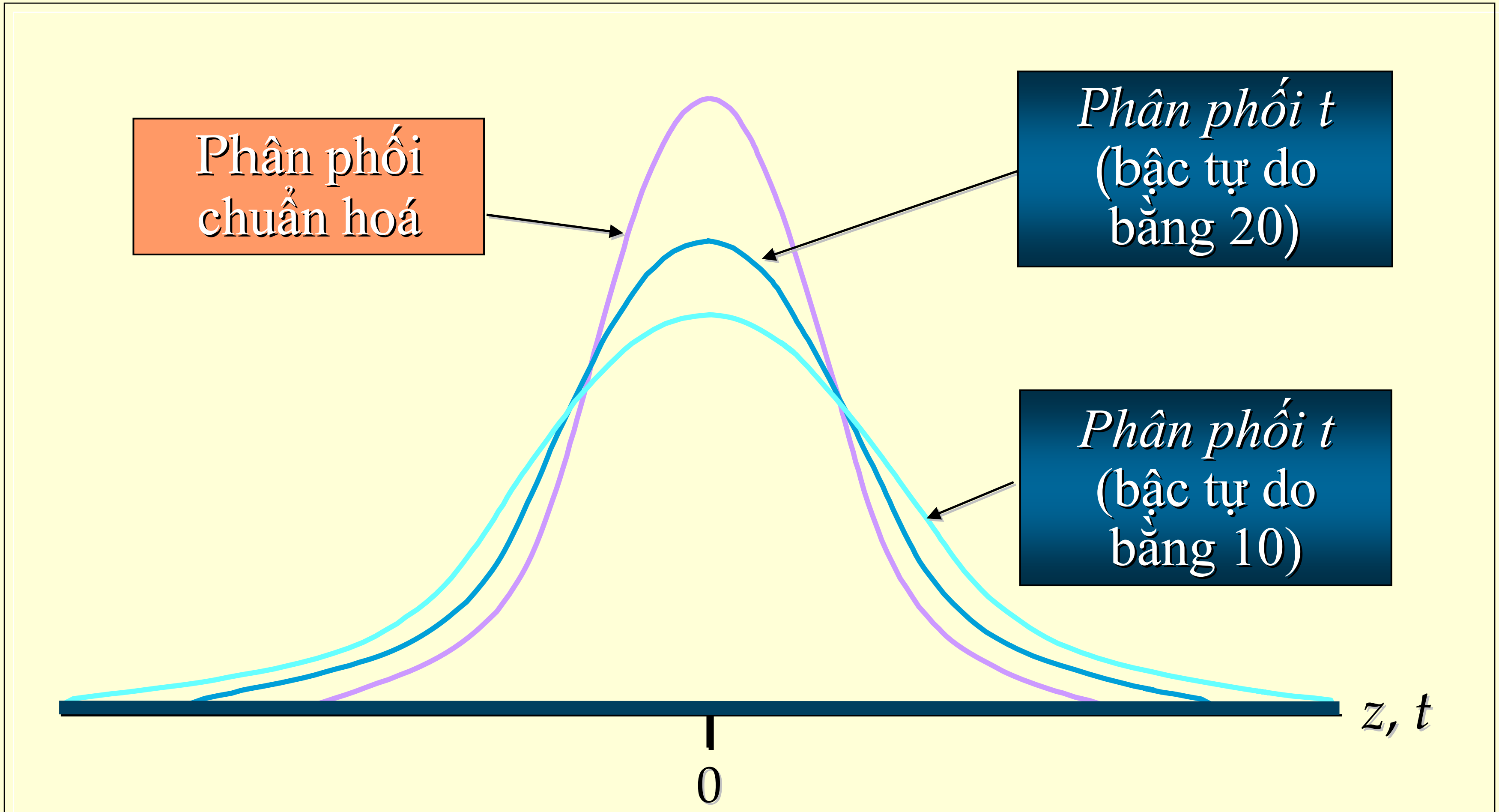
$$t_{(k)} = \frac{Z}{\sqrt{\chi^2_k / k}}$$

  tuân theo phân phối Student (phân phối T) với k bậc tự do.
- Phân phối T có dạng như phân phối chuẩn hoá, phân phối T có đuôi dày hơn so với phân phối chuẩn hoá, khi k tiến đến vô hạn thì phân phối T tiến dần đến phân phối chuẩn hoá.
- $\mu = 0$ và $\sigma^2 = k/(k-2) > 1$.
- Khi biết một biến ngẫu nhiên nào đó tuân theo phân phối t, chúng ta có thể trình bày xác suất như sau:

  $P(t(k,\alpha/2) \le tk \le t(k,1-\alpha/2))=1- \alpha$

Một số phân phối xác suất thường dùng

5. t-Distribution/ Phân phối T



Một số phân phối xác suất thường dùng

6. F-Distribution/ Phân phối F

- Phân phối F, là phân phối của tỉ lệ giữa hai biến ngẫu nhiên có phân phối chi-bình phương

$$F_{(k1,k2)} = \frac{\chi^2_{k1} / k_1}{\chi^2_{k2} / k_2}$$

- Phân phối F lệch về bên trái, khi bậc tự do k_1 và k_2 đủ lớn, phân phối F tiến đến phân phối chuẩn.
- $\mu = k_2/(k_2-2)$ với điều kiện $k_2>2$

- ***Lưu ý :*** *Khi bậc tự do đủ lớn thì các phân phối χ^2, phân phối T và phân phối F tiến đến phân phối chuẩn. Các phân phối này được gọi là phân phối có liên quan đến phân phối chuẩn*

2.5. Ước lượng thống kê

Ước lượng

Ước lượng (Estimator) và hàm ước lượng

- Là biến ngẫu nhiên hay các tham số thống kê của mẫu được dùng để ước lượng các tham số thống kê chưa biết của tập hợp chính.
- Ước lượng của tham số thống kê θ của tập hợp chính được ký hiệu là $\hat{\theta}$
- Dựa vào mẫu $\{x_1,x_2...,x_n\}$ người ta lập ra Hàm:

  $\hat{\theta} = \hat{\theta}(x_1,x_2,....,x_n)$ để ước lượng cho θ. được gọi là hàm ước lượng của θ hay gọi tắt là ước lượng của θ.
- $\hat{\theta}$ chỉ phụ thuộc vào giá trị quan sát x_1, x_2, ... ,x_n chứ không phụ thuộc vào các tham số chưa biết θ của tập hợp chính.

2.5. Ước lượng thống kê

Ước lượng điểm

Giá trị ước lượng (Estimate) hay còn gọi là giá trị ước lượng điểm

- Là giá trị cụ thể của ước lượng $\hat{\theta}$ và được xem như giá trị ước lượng của tham số thống kê θ của tập hợp chính.

Tham số thống kê và tập hợp chính (Population Parameter)	Ước lượng (Estimation)	Giá trị ước lượng Estimate (Point estimate)
Số trung bình	μ	$\bar{X}$
Phương sai	σ_x^2	S_x^2
Độ lệch chuẩn	σ_x	S_x
Tỷ lệ	p	$\hat{f}$

2.5. Ước lượng thống kê

Ước lượng điểm

Ước lượng không chệch:

- Ước lượng θ được gọi là ước lượng không chệch của tham số thống kê θ nếu kỳ vọng của $\hat{\theta}$ là θ

$$E(\hat{\theta}) = \theta$$

Thí dụ

$E(\overline{X}) = \mu$ => $\overline{X}$ là ước lượng không chệch của μ

$E(S_x^2) = \sigma_x^2$ => S_x^2 là ước lượng không chệch cuả σ_x^2

$E(\hat{f}) = p$ => $\hat{f}$ là ước lượng không chệch của p

2.5. Ước lượng thống kê

Ước lượng điểm

Ước lượng hiệu quả tốt nhất:

- Gọi $\hat{\theta}_1$ và $\hat{\theta}_2$ là 2 ước lượng không chệch của θ dựa trên số lượng của mẫu quan sát giống nhau

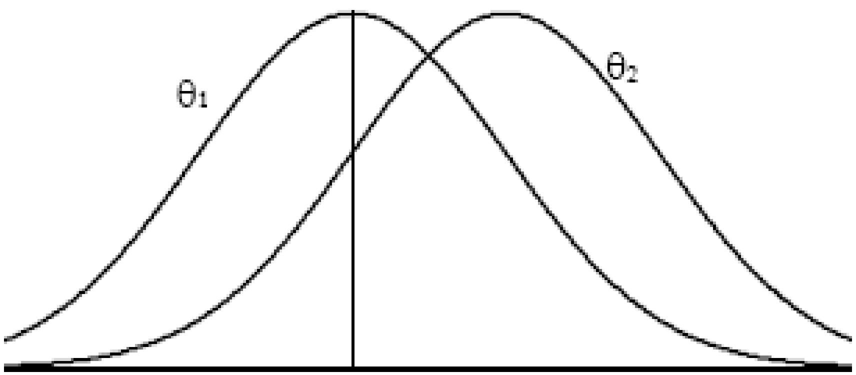
  $\hat{\theta}_1$ được gọi là hiệu quả hơn $\hat{\theta}_2$ nếu: Var ($\hat{\theta}_1$) < Var ($\hat{\theta}_2$)

$$\text{Hiệu quả tương đối} = \frac{Var(\hat{\theta}_2)}{Var(\hat{\theta}_1)}$$

- Nếu $\hat{\theta}$ là ước lượng không chệch của θ và nếu không có một ước lượng không chệch nào có phương sai nhỏ hơn phương sai của $\hat{\theta}$ thì $\hat{\theta}$ được gọi là ước lượng tốt nhất (Best Estimator) hay $\hat{\theta}$ còn gọi là ước lượng không chệch có phương sai nhỏ nhất của θ (Minimum Variance Unbiased Estimator of θ)

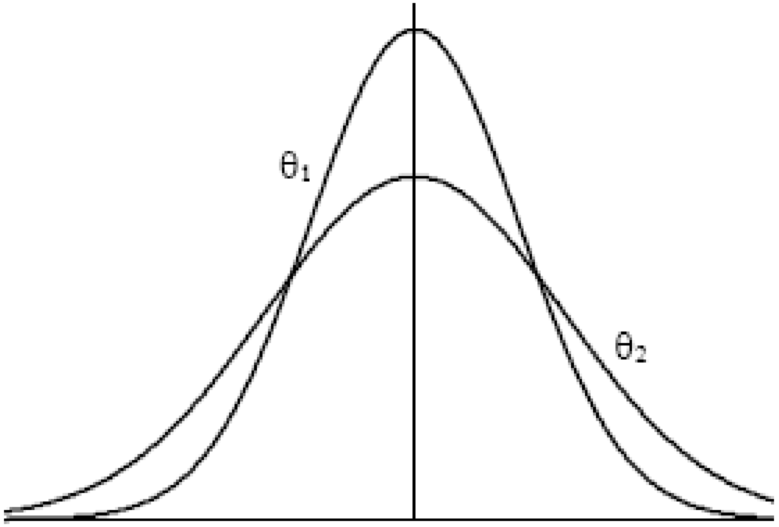
2.5. Ước lượng thống kê

Ước lượng điểm



$\hat{\theta}_1$: ước lượng không chệch của θ

$\hat{\theta}_2$: ước lượng chệch của θ



$\hat{\theta}_1 \hat{\theta}_2$: ước lượng không chệch của θ

$\hat{\theta}_1$ ước lượng hiệu quả hơn $\hat{\theta}_2$:

2.5. Ước lượng thống kê

Khoảng tin cậy

a) Ước lượng khoảng và giá trị ước lượng khoảng (Interval Estimator And Interval Estimate).

- Ước lượng khoảng: Ước lượng khoảng đối với tham số thống kê của tập hợp chính θ là một quy tắc dựa trên thông tin của mẫu để xác định miền (Range) hay khoảng (Interval) mà tham số θ hầu như nằm trong đó.
- Gía trị ước lượng khoảng: là giá trị cụ thể của miền hay khoảng mà tham số θ nằm trong đó.

b) Khoảng tin cậy và độ tin cậy (Confidence Interval and Level of Confidence)

- Gọi θ là tham số thống kê chưa biết. Giả sử dựa trên thông tin của mẫu ta có thể xác định được 2 biến ngẫu nhiên A và B sao cho

  $P(A < \theta < B) = 1 - \alpha$ với $0 < \alpha < 1$
- Nếu giá trị cụ thể của biến ngẫu nhiên A và B là a và b thì khoảng (a,b) từ a đến b được gọi là khoảng tin cậy của θ với xác suất là $(1 - \alpha)$
- Xác suất $(1 - \alpha)$ được gọi là độ tin cậy của khoảng.
- ***Ghi chú:***
  - *Trong thực tế, độ tin cậy (1-α) do nhà thống kê chọn theo yêu cầu của mình, thông thường độ tin cậy được chọn là 0,90; 0,95; 0,99...*
  - *α là xác suất sai lầm khi chọn khoảng tin cậy (a, b)*

2.5. Ước lượng thống kê

Khoảng tin cậy

c) Ước lượng khoảng cho kỳ vọng tham số μ trong phân phối chuẩn N(μ, σ²)

Cho mẫu ngẫu nhiên (X1, X2, ..., Xn) của ĐLNN X có phân phối chuẩn N(μ, σ^2), khoảng ước lượng của tham số μ được tính như sau:

- Trường hợp σ^2 đã biết: Khoảng ước lượng của tham số μ với độ tin cậy 1 - α là

$$\overline{X} - x_\alpha \frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + x_\alpha \frac{\sigma}{\sqrt{n}}$$

Trong đó: x_α là số được tra từ bảng phân phối chuẩn tắc N(0, 1) sao cho $F(x_\alpha) = 1 - \alpha/2$ (sử dụng hàm MS-Excel: x_a = NORMSINV(1 - a/2)).

2.5. Ước lượng thống kê

Khoảng tin cậy

- Trường hợp σ^2 chưa biết: Khoảng ước lượng của tham số μ với độ tin cậy $1 - \alpha$ là

$$\overline{X} - t_\alpha \frac{S_n^*(X)}{\sqrt{n}} < \mu < \overline{X} + t_\alpha \frac{S_n^*(X)}{\sqrt{n}}$$

$$S_n^{*2}(X) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

Trong đó: nếu $n \geq 30$ thì ta tra giống x_α ở trên; nếu $n < 30$ thì ta tra trong bảng phân phối Student với $n - 1$ bậc tự do (bảng 2 phía) và mức ý nghĩa α, (sử dụng hàm MS-Excel: t_α = TINV(α , n -1))

Khoảng tin cậy

- Bảng tra phân phối T

	Giới hạn trên		
df	.25	.10	**.05**
1	1.000	3.078	6.314
2	0.817	1.886	2.920
3	0.765	1.638	2.353

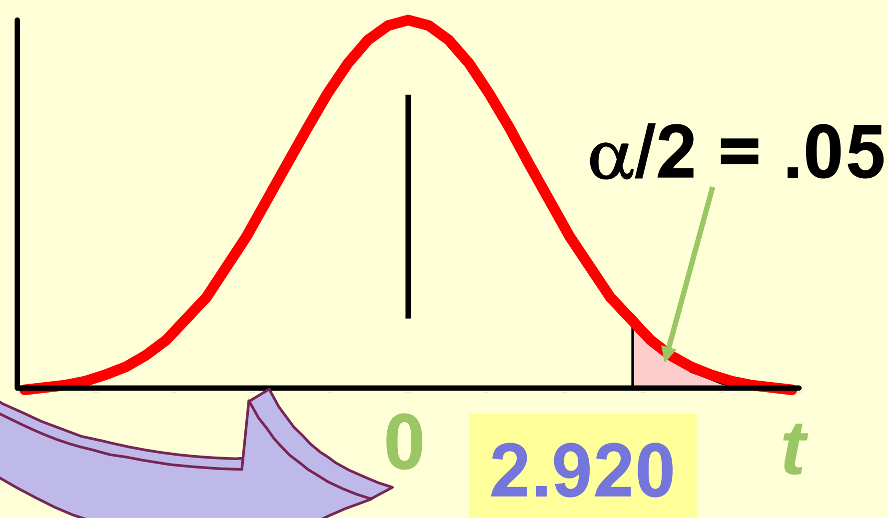
t Values

Khi: $n = 3$

$df = n - 1 = 2$

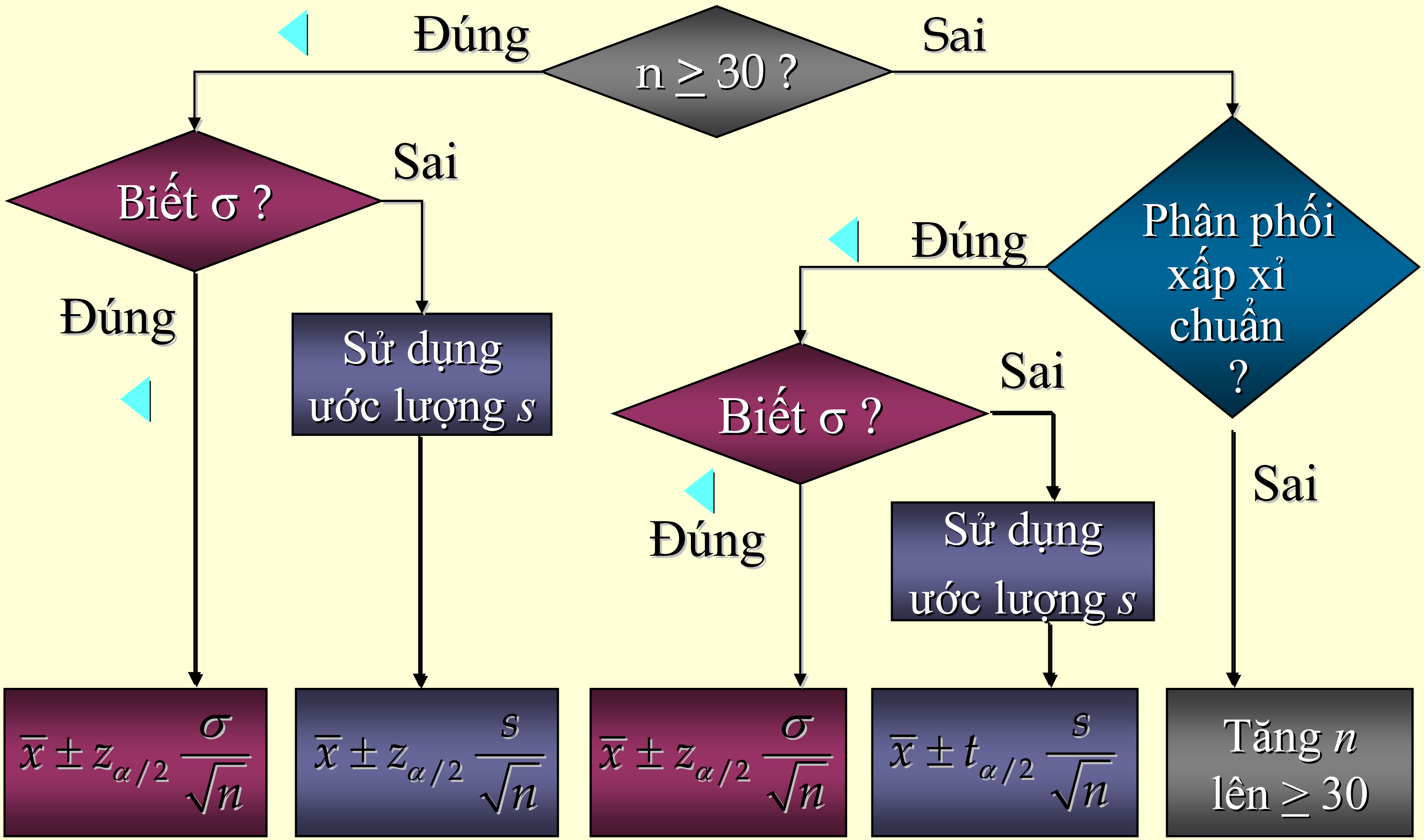
$\alpha = .10$

$\alpha/2 = .05$



2.5. Ước lượng thống kê

Tổng hợp các phân phối sử dụng trong ước lượng



2.5. Ước lượng thống kê

Khoảng tin cậy- Ví dụ

- **Ví dụ 1**: Tìm khoảng ước lượng của kỳ vọng μ với độ tin cậy 0,95 của ĐLNN X có phân phối chuẩn nếu biết trung bình mẫu là 14, độ lệch bình phương trung bình là 5 và kích thước mẫu là 25.

  Giải:

  Trường hợp này cho biết độ lệch bình phương trung bình là 5 tức là biết phương sai, ta có:

$$\alpha = 0,05 \;\Rightarrow F(x_\alpha) = 1 - \frac{\alpha}{2} = 0,975 \;\Rightarrow x_\alpha = NORMSINV(0,975) = 1,96$$

$$\overline{X} - x_\alpha \frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + x_\alpha \frac{\sigma}{\sqrt{n}} \;\Rightarrow 14 - 1,96\frac{5}{\sqrt{25}} < \mu < 14 + 1,96\frac{5}{\sqrt{25}}$$

$$12,4 < \mu < 15,96$$

2.5. Ước lượng thống kê

Khoảng tin cậy- Ví dụ

- **Ví dụ 2**: Tìm khoảng ước lượng của kỳ vọng µ với độ tin cậy 0,95 của ĐLNN X có phân phối chuẩn, kích thước mẫu là 25 và giả sử tìm được trung bình mẫu là 14, phương sai mẫu điều chỉnh là 9 .

  Giải:

  Trường hợp này chỉ biết phương sai mẫu điều chỉnh là 9 tức là không biết phương sai, ta có:

$$\alpha = 0{,}05;\ n = 25;\ S^{*2}(X) = 9 \Rightarrow S^{*}(X) = 3;\quad t_{\alpha} = TINV(0{,}05{,}24) = 2{,}06$$

$$\overline{X} - t_{\alpha}\frac{S^{*}}{\sqrt{n}} < \mu < \overline{X} - t_{\alpha}\frac{S^{*}}{\sqrt{n}} \Rightarrow 14 - 2{,}06\frac{3}{\sqrt{25}} < \mu < 14 + 2{,}06\frac{3}{\sqrt{25}}$$

$$12{,}764 < \mu < 15{,}236$$

2.5. Ước lượng thống kê

Khoảng tin cậy- Bài tập

1. Điểm trung bình môn toán của 100 thí sinh dự thi vào ĐHBK là 6 với độ lệch mẫu điều chỉnh là 2,5. Hãy ước lượng điểm trung bình môn toán của toàn thể thí sinh với độ tin cậy 95%.
2. Tuổi thọ của 1 loại bóng đèn được biết theo quy luật chuẩn với độ lệch chuẩn 100 giờ. Chọn ngẫu nhiên 100 bóng để thử nghiệm, thấy mỗi bóng tuổi thọ trung bình là 1000 giờ. Hãy ước lượng tuổi thọ trung bình của bóng đèn xí nghiệp A sản xuất với độ tin cậy 95%.
3. Điều tra năng suất lúa trên diện tích 100ha ở một vùng ta có bảng số liệu sau:

Năng suất (tấn)	21	24	25	26	28	32	34
Diện tích (ha)	10	20	30	15	10	10	5

Hãy ước lượng năng suất lúa trung bình của vùng đó với độ tin cậy 95%.

2.6. Kiểm định giả thiết thống kê

Nguyên lý cơ bản

- Giả thuyết thống kê là một giả sử hay một phát biểu có thể đúng, có thể sai liên quan đến tham số, luật phân phối hay tính chất của biến ngẫu nhiên. Khi thực hiện kiểm định, người ta thiết lập cặp giả thiết thống kê, **Giả thuyết không** và **giả thuyết ngược lại** (giả thiết đối).
  - *Giả thuyết không: là sự giả sử mà chúng ta muốn kiểm định thường được ký hiệu là H_0.*
  - *Giả thuyết ngược lại: Việc bác bỏ giả thuyết không sẽ dẫn đến việc chấp nhận giả thuyết ngược lại. Giả thuyết ngược lại thường được ký hiệu là H_1.*

    Ví dụ:

    $H_0 : \mu = 0.5$

    $H_1 : \mu \neq 0.5$

2.6. Kiểm định giả thiết thống kê

Nguyên lý cơ bản

- Tất cả các giá trị có thể có của các đại lượng thống kê trong kiểm định có thể chia làm 2 miền: miền bác bỏ và miền chấp nhận. Giá trị chia đôi hai miền được gọi là giá trị giới hạn (Critical value)
  - *Miền bác bỏ là miền chứa các giá trị làm cho giả thuyết Ho bị bác bỏ.*
  - *Miền chấp nhận là miền chứa các giá trị giúp cho giả thuyết Ho không bị bác bỏ.*
- Giả thiết không và giả thiết đối có thể là giả thiết đơn hay giả thiết kép. Một giả thiết được gọi là đơn nếu nó đưa ra 1 giá trị cụ thể cho tham số (ví dụ H_0:β=0.5). Một giả thiết được gọi là kép nếu nó đưa ra một khoảng giá trị của phân bố xác suất (ví dụ H_0: $\beta > 0.5$). Liên quan đến vấn đề này người ta có kiểm định hai phía và kiểm định một phía.

2.6. Kiểm định giả thiết thống kê

Nguyên lý cơ bản

- Việc kiểm định được thực hiện theo các bước như sau:
  - B1: Lập 1 mẫu ngẫu nhiên (X1, X2, X3, …, Xn) cho bnn X.
  - B2: Tìm một hàm G = f((X1, X2, X3, …, Xn, Z), sao cho luật phân bố của hàm G đã biết. (Z là thông số liên quan đến giả thiết cần kiểm định).
  - B3: Tìm một miền Wα sao cho xác suất để giá trị của hàm G rơi vào miền này đúng bằng α với 0<α<1 và đủ bé để sao cho trong một phép thử rất khó có thể thu được giá trị hàm G rơi vào miền Wα.
  - B4: Lấy một mẫu cụ thể (X1, X2, X3, …, Xn) tính giá trị của hàm G cho mẫu này: $g_0 = G_0$ (X1, X2, X3, …, Xn)
  - Khi đó có các trường hợp:
    - *TH1: g_0 thuộc Wα => bác bỏ giả thiết H_0 ở mức ý nghĩa α*
    - *TH2: g_0 không thuộc Wα => không có cơ sở bác bỏ giả thiết H_0 (chấp nhận).*

2.6. Kiểm định giả thiết thống kê

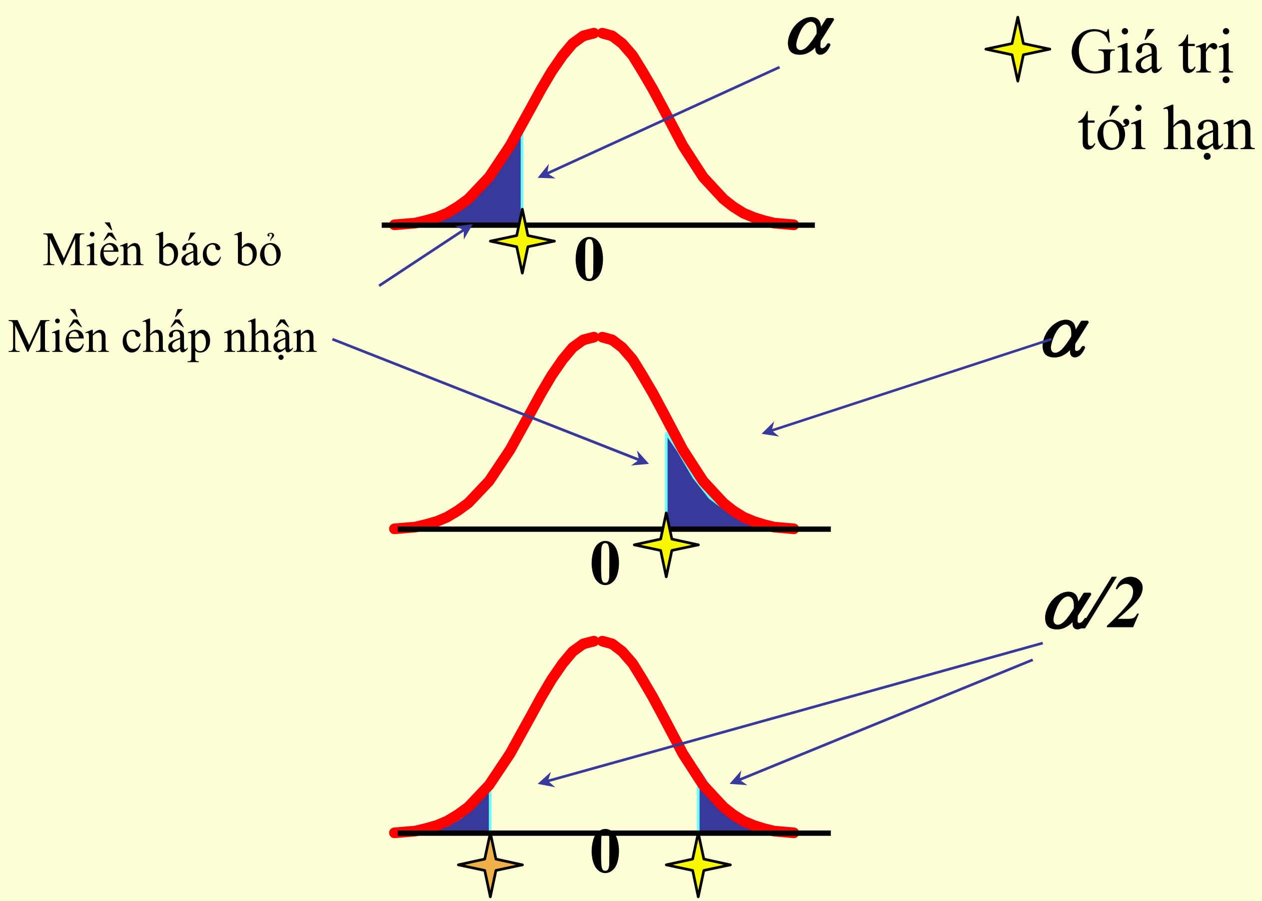
Nguyên lý cơ bản

- Miền bác bỏ và miền chấp nhận H_0

$$H_0: \mu \geq 3.5$$
$$H_1: \mu < 3.5$$

$$H_0: \mu \leq 3.5$$
$$H_1: \mu > 3.5$$

$$H_0: \mu = 3.5$$
$$H_1: \mu \neq 3.5$$



2.6. Kiểm định giả thiết thống kê

Các kiểm định thông dụng

- Kiểm định 1 phía cho trung bình của tổng thể
  - Giả định:
    - *Tổng thể có phân phối chuẩn.*
    - *Giả thiết không là ≤ hoặc ≥*
    - *Phương sai đã biết (σ^2 đã biết)*
  - Thống kê kiểm định: sử dụng phân phối Z

$$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$$

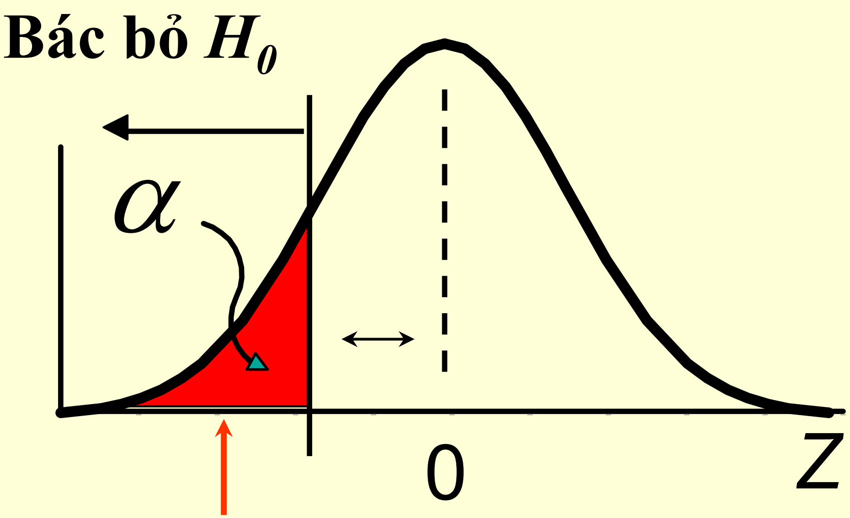
2.6. Kiểm định giả thiết thống kê

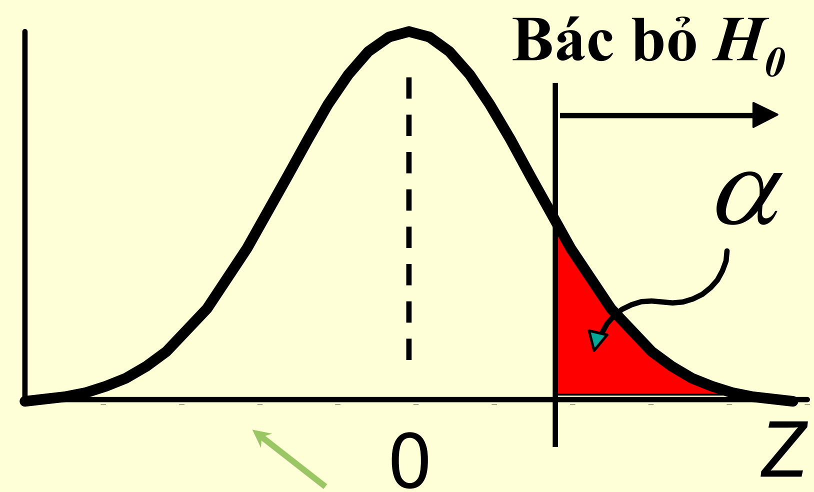
Các kiểm định thông dụng

- Kiểm định 1 phía cho trung bình của tổng thể

$$H_0: \ \mu \geq \mu_0$$
$$H_1: \ \mu < \mu_0$$

$$H_0: \ \mu \leq \mu_0$$
$$H_1: \ \mu > \mu_0$$





Miền bác bỏ

Miền chấp nhận

2.6. Kiểm định giả thiết thống kê

Các kiểm định thông dụng

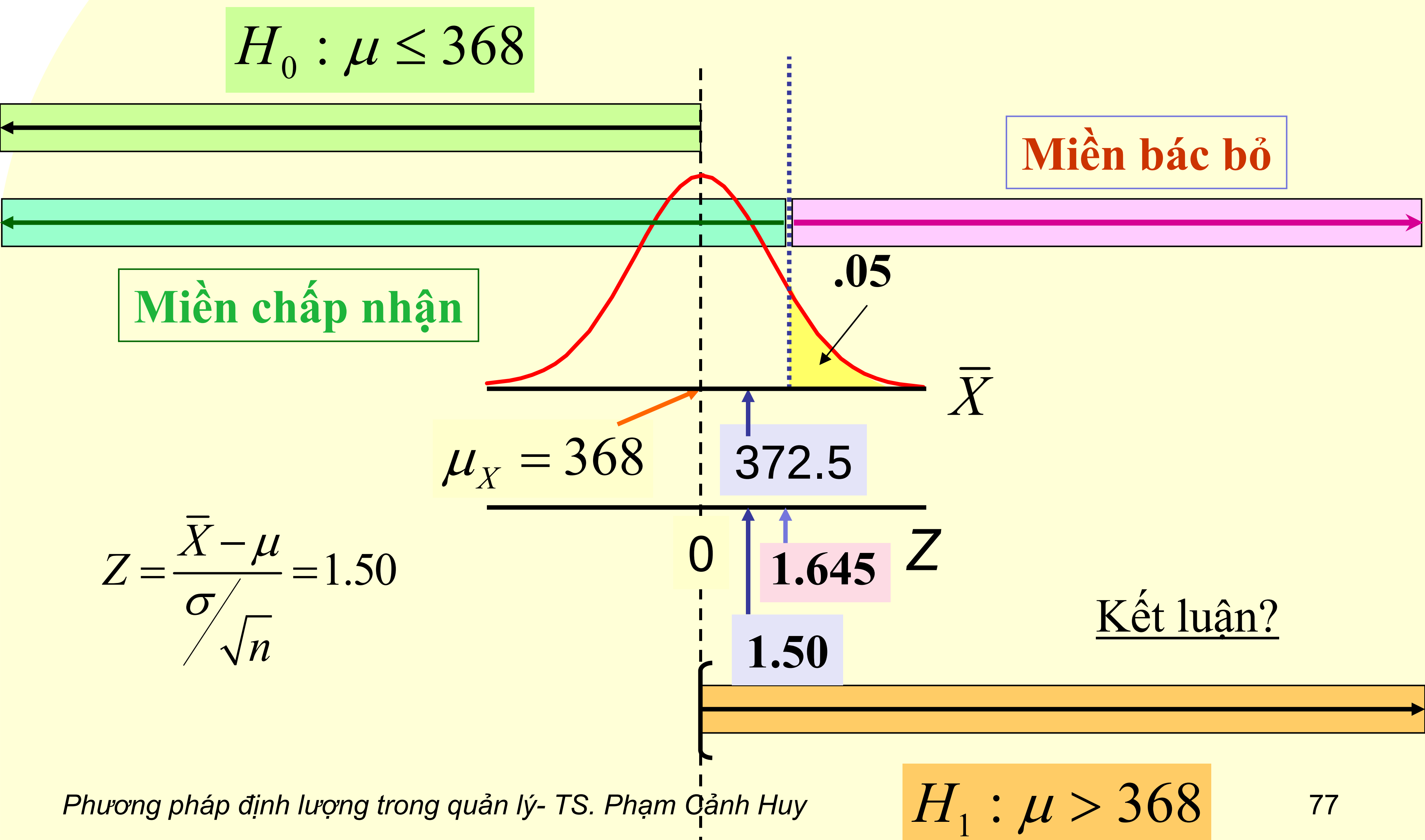
- Kiểm định 1 phía cho trung bình của tổng thể: Ví dụ

Để kiểm tra xem trọng lượng trung bình của hộp ngũ cốc có nhiều hơn 368 grams hay không? Người ta lấy mẫu 25 hộp và thấy rằng trọng lượng trung bình bằng 372.5. Công ty xác định độ lệch chuẩn cho phép là σ = 15 grams. Hãy thực hiện kiểm định với α = 0.05

$$H_0: \ \mu \le 368$$
$$H_1: \ \mu > 368$$

2.6. Kiểm định giả thiết thống kê

Các kiểm định thông dụng



2.6. Kiểm định giả thiết thống kê

Các kiểm định thông dụng

- Kiểm định 2 phía cho trung bình của tổng thể: Ví dụ

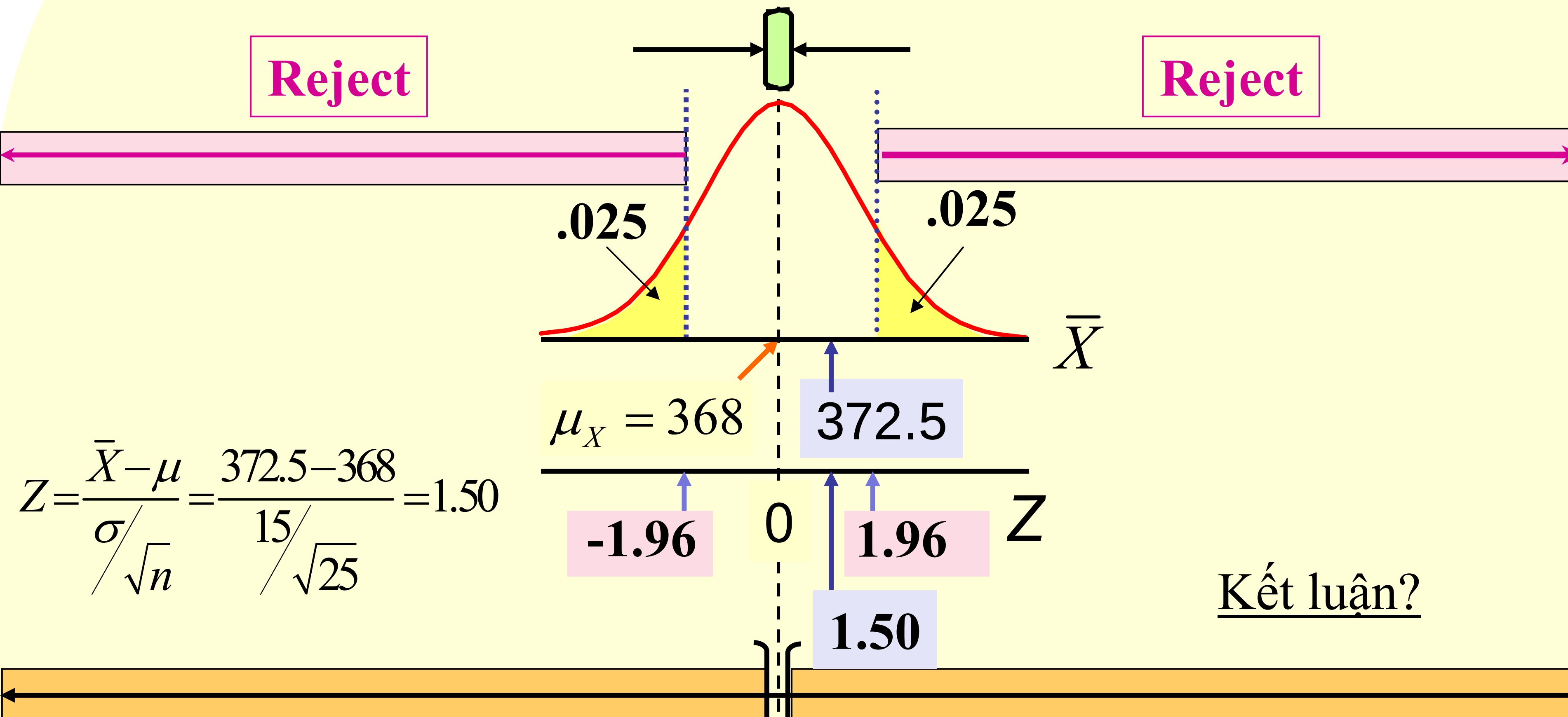
Để kiểm tra xem trọng lượng trung bình của hộp ngũ cốc có bằng 368 grams hay không? Người ta lấy mẫu 25 hộp và thấy rằng trọng lượng trung bình bằng 372.5. Công ty xác định độ lệch chuẩn cho phép là σ = 15 grams. Hãy thực hiện kiểm định với α = 0.05

$$H_0: \mu = 368$$
$$H_1: \mu \neq 368$$

2.6. Kiểm định giả thiết thống kê

Các kiểm định thông dụng

$$H_0 : \mu = 368$$



$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} = \frac{372.5 - 368}{15 / \sqrt{25}} = 1.50$$

Kết luận?

$$H_1 : \mu \neq 368$$

2.6. Kiểm định giả thiết thống kê

Các kiểm định thông dụng

- Trường hợp kiểm định cho trung bình của tổng thể khi chưa biết phương sai (hay độ lệch chuẩn), ta sẽ sử dụng thống kê T.
- Việc kiểm định cho phương sai của tổng thể ta sẽ sử dụng thống kê χ^2 và thực hiện tương tự.

Chương 3

PHÂN TÍCH HỒI QUI

Nội dung

3.1. Khái niệm phân tích hồi qui

3.2. Mô hình hồi qui đơn biến

3.3. Mô hình hồi qui đa biến

3.1. Khái niệm phân tích hồi qui

Khái niệm phân tích hồi qui

- Phân tích hồi quy là tìm quan hệ phụ thuộc của một biến, được gọi là biến phụ thuộc vào một hoặc nhiều biến khác, được gọi là biến độc lập nhằm mục đích ước lượng hoặc tiên đoán giá trị kỳ vọng của biến phụ thuộc khi biết trước giá trị của biến độc lập.
- *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. 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*

3.1. Khái niệm phân tích hồi qui

Hồi qui tổng thể và hồi qui mẫu

- Hàm hồi quy tổng thể (PRF):
  - $E(Y/X=X_i) = \beta_1 + \beta_2 X$
  - Đối với một quan sát cụ thể thì giá trị biến phụ thuộc lệch khỏi kỳ vọng toán, vậy:

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

  Trong đó:

  - *β_1 và β_2 là các tham số của mô hình*
  - *ui là Sai số của hồi quy hay còn được gọi là nhiễu ngẫu nhiên. Nhiễu ngẫu nhiên hình thành có thể do: Bỏ sót biến giải thích, Sai số khi đo lường biến phụ thuộc, Các tác động không tiên đoán được hay Dạng hàm hồi quy không phù hợp.*

3.1. Khái niệm phân tích hồi qui

Hồi qui tổng thể và hồi qui mẫu

- Hàm hồi quy mẫu (SRF):
  - Trong thực tế hiếm khi chúng có số liệu của tổng thể mà chỉ có số liệu mẫu. Chúng ta phải sử dụng dữ liệu mẫu để ước lượng hàm hồi quy tổng thể.
  - Hàm hồi quy mẫu được biểu diễn:

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

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* β_1
- $\hat{\beta}_2$ *là ước lượng của* β_2

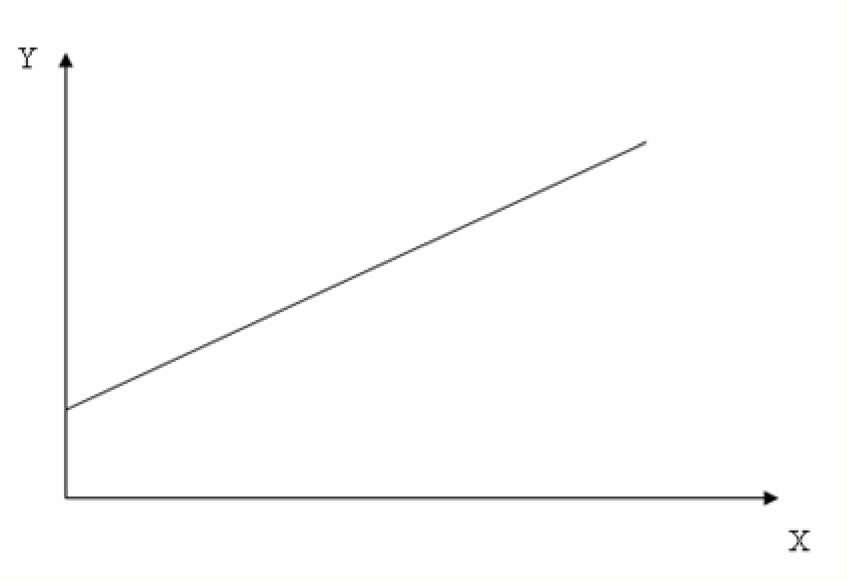
3.1. Khái niệm phân tích hồi qui

Một số dạng hàm cơ bản trong phân tích hồi qui

- ***Dạng Hàm Tuyến tính*** :

Dạng hàm này có phương trình:

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



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 β_2 đơn vị. 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.

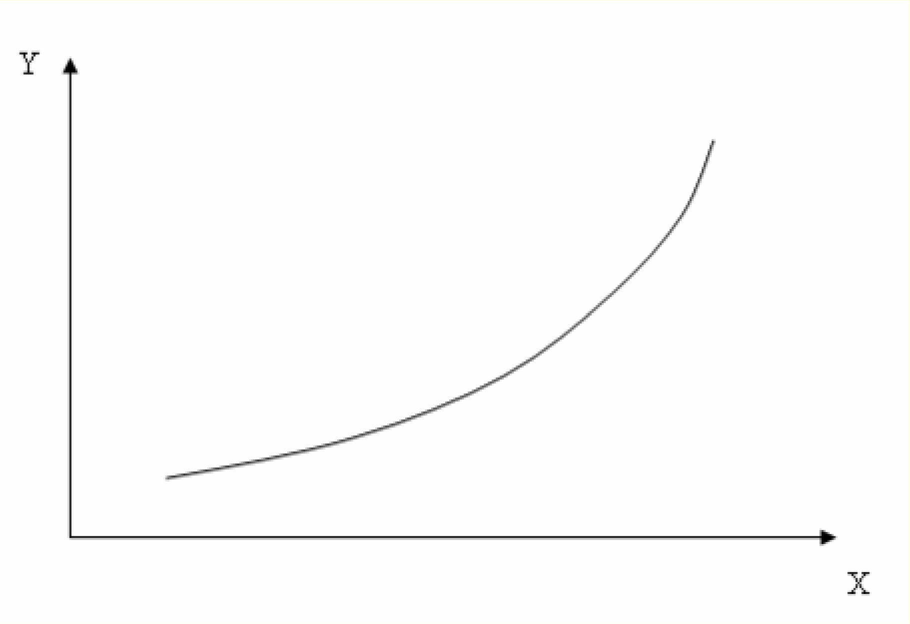
3.1. Khái niệm phân tích hồi qui

Một số dạng hàm cơ bản trong phân tích hồi qui

- ***Dạng Hàm Bậc hai*** :

Dạng hàm này có phương trình:

$$Y_i = \beta_1 + \beta_2 X_i + \beta_3 X_i^2 + u_i$$



Khi *X* tăng thêm một đơn vị thì *Y* tăng thêm $\beta_2 + 2\beta_3 X_i$ đơn vị. Nếu $\beta_3 > 0$, thì khi *X* tăng lên, tác động bổ sung của *X* đến *Y* cũng tăng lên; nếu $\beta_3 < 0$, thì khi *X* tăng lên, tác động bổ sung của *X* đến *Y* giảm xuống. Nếu bạn có đường biểu diễn chi phí thì chi phí biên (tức là số đơn vị mà *C* tăng lên khi *Q* tăng lên thêm một đơn vị) sẽ là $MC = \beta_2 + 2\beta_3 Q$

3.1. Khái niệm phân tích hồi qui

Một số dạng hàm cơ bản trong phân tích hồi qui

- ***Dạng Hàm Logarít*** :

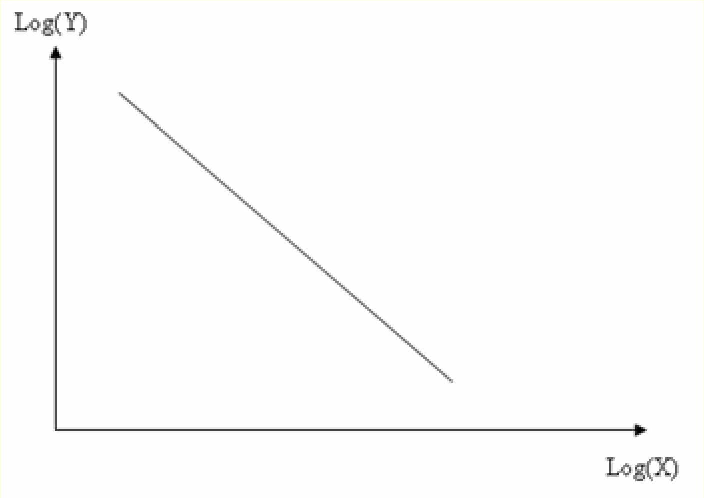
Dạng hàm này có phương trình:

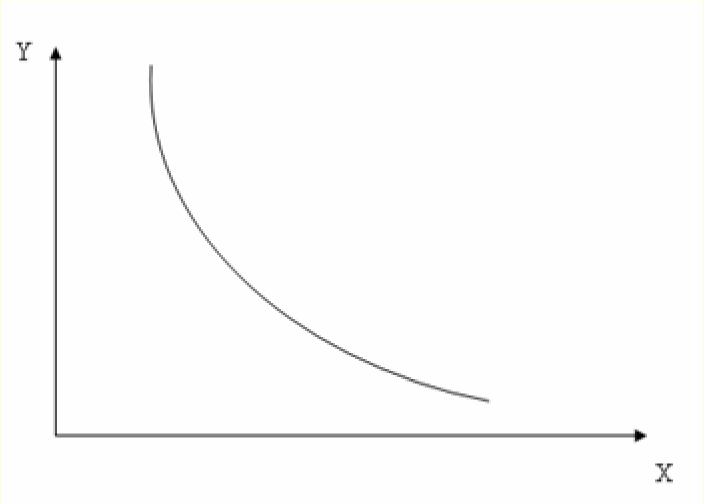
$$\ln Y_i = \beta_1 + \beta_2 \ln X_i + u_i$$

Giải thích dạng hàm này là nếu *X* thay đổi 1% thì *Y* sẽ thay đổi β_2%; đây là tính chất đặc biệt của quan hệ logarít.

Hàm Cobb-Douglas có dạng:

$$Y = \beta_1 K^{\beta 2} L^{\beta 3} e^{u}$$





3.1. Khái niệm phân tích hồi qui

Một số dạng hàm cơ bản trong phân tích hồi qui

- ***Dạng Hàm bán-lôgarít (Semilog)***:

Dạng hàm này có phương trình:

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

Nếu *X* tăng thêm 1 đơn vị thì *Y* tăng thêm [β_2*100] %. Một số ứng dụng hữu ích cho dạng hàm này. Ví dụ, quan hệ giữa tiền lương và trình độ giáo dục.

Dạng hàm này có phương trình:

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

Khi *X* tăng 1%, thì *Y* tăng [β_2/100] đơn vị.

3.1. Khái niệm phân tích hồi qui

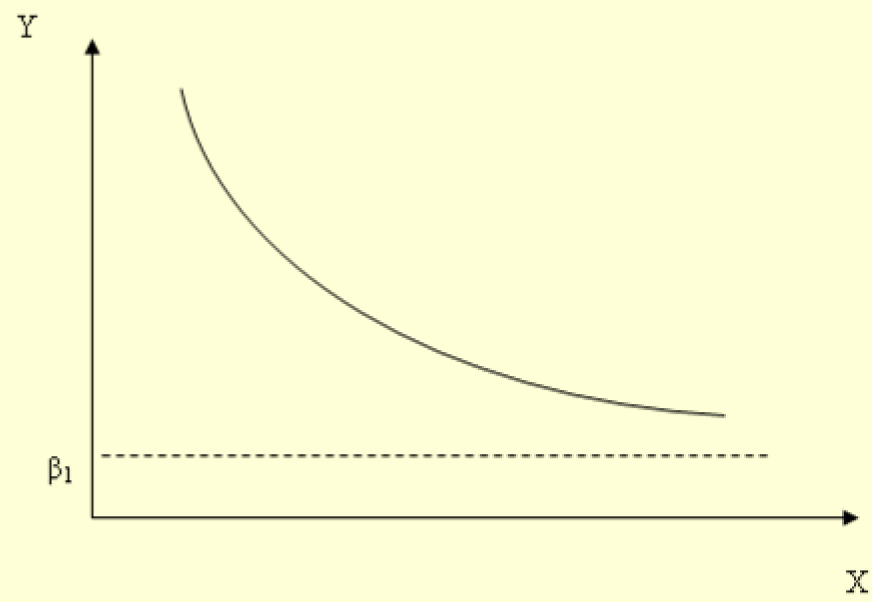
Một số dạng hàm cơ bản trong phân tích hồi qui

- ***Dạng Hàm Nghịch đảo:***

Dạng hàm này có phương trình:

$Y_i = \beta_1 + \beta_2(1/X_i) + u_i$

Dạng hàm nghịch đảo thường được sử dụng khi *Y* và *X* đều dương và khi đường biểu diễn quan hệ giữa chúng dốc xuống (β_1>0 và β_2>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.



3.1. Khái niệm phân tích hồi qui

Một số dạng hàm cơ bản trong phân tích hồi qui

- ***Tổng quát :***

Ta cũng có thể kết hợp vài dạng hàm khác nhau trong một hồi qui, ví dụ:

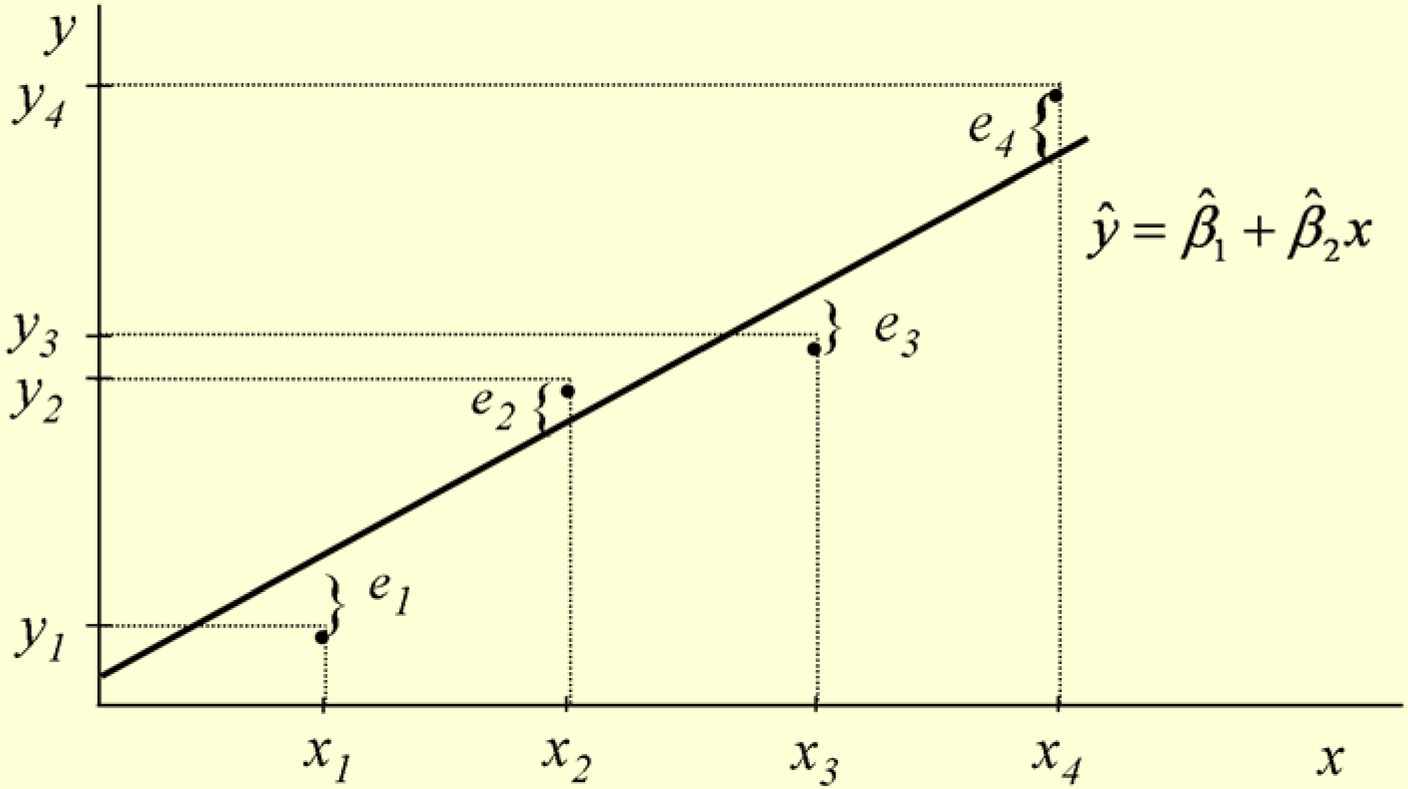
$$Y_i = \beta_1 + \beta_2(1/X_2) + \beta_3 \ln X_3 + \beta_4 X_4 + \beta_5 X_4^2 + u_i$$

nếu ta làm thế, ta thường phải có các lý do thỏa đáng để nghĩ rằng hình dạng của quan hệ giữa *Y* và X_2 là khác với các hình dạng của quan hệ giữa *Y* và X_3 , và *Y* và X_4.

3.2. Mô hình hồi qui đơn

Phương pháp bình phương nhỏ nhất

- Đây là phương pháp được đưa ra bởi nhà toán học Đức Carl Friedrich Gauss, ký hiệu OLS (ordinary least squares). Tư tưởng của phương pháp này là cực tiểu tổng bình phương các phần dư.



OLS- Hồi qui đơn

- Hàm hồi qui mẫu: $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 x_i$

Đặt $L = \sum_i (y_i - \hat{y}_i)^2 = \sum_i (y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i)^2$

Ta thấy rằng các tham số qui mẫu sẽ là nghiệm của hệ thống phương trình sau

$$\frac{\partial L}{\partial \hat{\beta}_1} = -2\sum_i (y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i) = 0 \qquad (1)$$

$$\frac{\partial L}{\partial \hat{\beta}_2} = -2\sum_i x_i (y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i) = 0 \qquad (2)$$

Từ (1) => $\sum_i (y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i) = 0 \Leftrightarrow \sum_i y_i - n\hat{\beta}_1 - \hat{\beta}_2 \sum_i x_i = 0$

OLS- Hồi qui đơn

- Đặt $\sum y_i = n\overline{y}$ $\sum x_i = n\overline{x}$

  Do vậy ta có thể viết: $n\overline{y} - n\hat{\beta}_1 - n\hat{\beta}_2\overline{x} = 0$ hay $\overline{y} - \hat{\beta}_1 - \hat{\beta}_2\overline{x} = 0$

  Từ (2) => $\sum_i x_i(y_i - \hat{\beta}_1 - \hat{\beta}_2 x_i) = 0$

  Do vậy: $\sum_i x_i(y_i - \overline{y} + \hat{\beta}_2\overline{x} - \hat{\beta}_2 x_i) = 0$

  $$\sum_i x_i y_i - \overline{y}\sum x_i + \hat{\beta}_2\overline{x}\sum x_i - \hat{\beta}_2\sum {x_i}^2 = 0$$

  $$\sum_i x_i y_i - n\overline{y}\overline{x} + \hat{\beta}_2 n\overline{x}^2 - \hat{\beta}_2\sum {x_i}^2 = 0$$

3.2. Mô hình hồi qui đơn

OLS- Hồi qui đơn

- Suy ra:

$$\hat{\beta}_2 = \frac{\sum x_i y_i - n\overline{xy}}{\sum x_i^2 - n\overline{x}^2} = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sum (x_i - \overline{x})^2} \quad \& \quad \hat{\beta}_1 = \overline{y} - \hat{\beta}_2 \overline{x}$$

- $\hat{\beta}_1, \hat{\beta}_2$ là các ước lượng của β_1 và β_2 được tính bằng phương pháp bình phương nhỏ nhất- được gọi là các ước lượng bình phương nhỏ nhất

3.2. Mô hình hồi qui đơn

OLS- Hồi qui đơn

Các giả thiết

- *Phương pháp bình phương nhỏ nhất (OLS) là phương pháp rất đáng tin cậy trong việc ước lượng các tham số của mô hình, tuy nhiên mô hình ước lượng phải thoả mãn các giả thiết. Khi thoả mãn các giả thiết, ước lượng bình phương nhỏ nhất (OLS) là ước lượng tuyến tính không chệch có hiệu quả nhất trong các ước lượng. Vì thế phương pháp OLS đưa ra* ***Ước Lượng Không chệch Tuyến Tính Tốt Nhất (BLUE).*** *Kết quả này được gọi là* ***Định lý Gauss–Markov****, Các giả thiết như sau.*

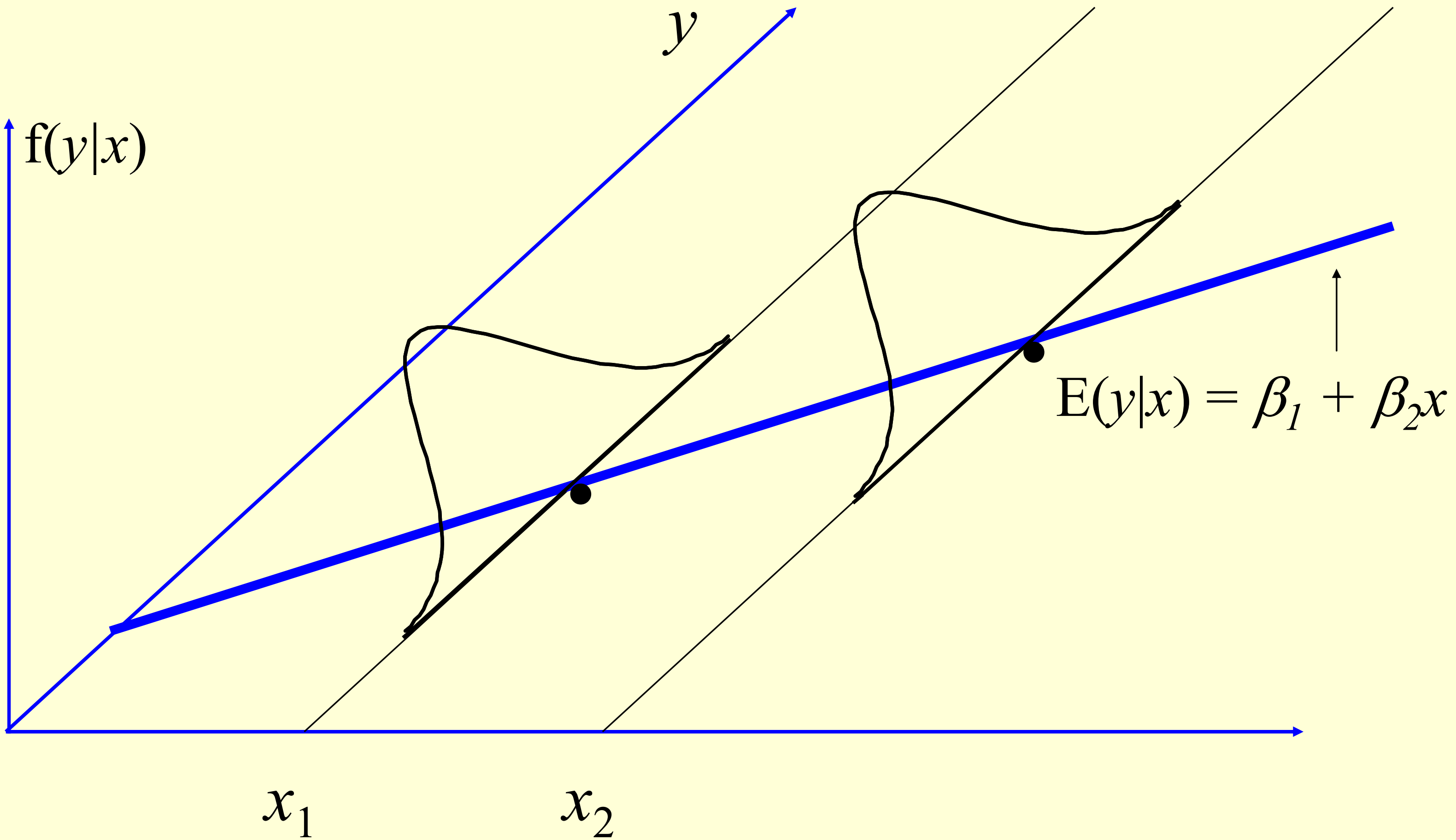
- Giả thiết

1. $E(u_i) = 0$	Kỳ vọng của các yếu tố ngẫu nhiên u_i bằng 0
2. $Var(u_i) = \sigma^2$	Phương sai bằng nhau với mọi u_i
3. $Cov(u_i, u_j) = 0$	Không có sự tương quan giữa các u_i
4. $Cov(u_i, x_i) = 0$	U và X không tương quan với nhau
5. u_i	Phân phối chuẩn

3.2. Mô hình hồi qui đơn

OLS- Hồi qui đơn

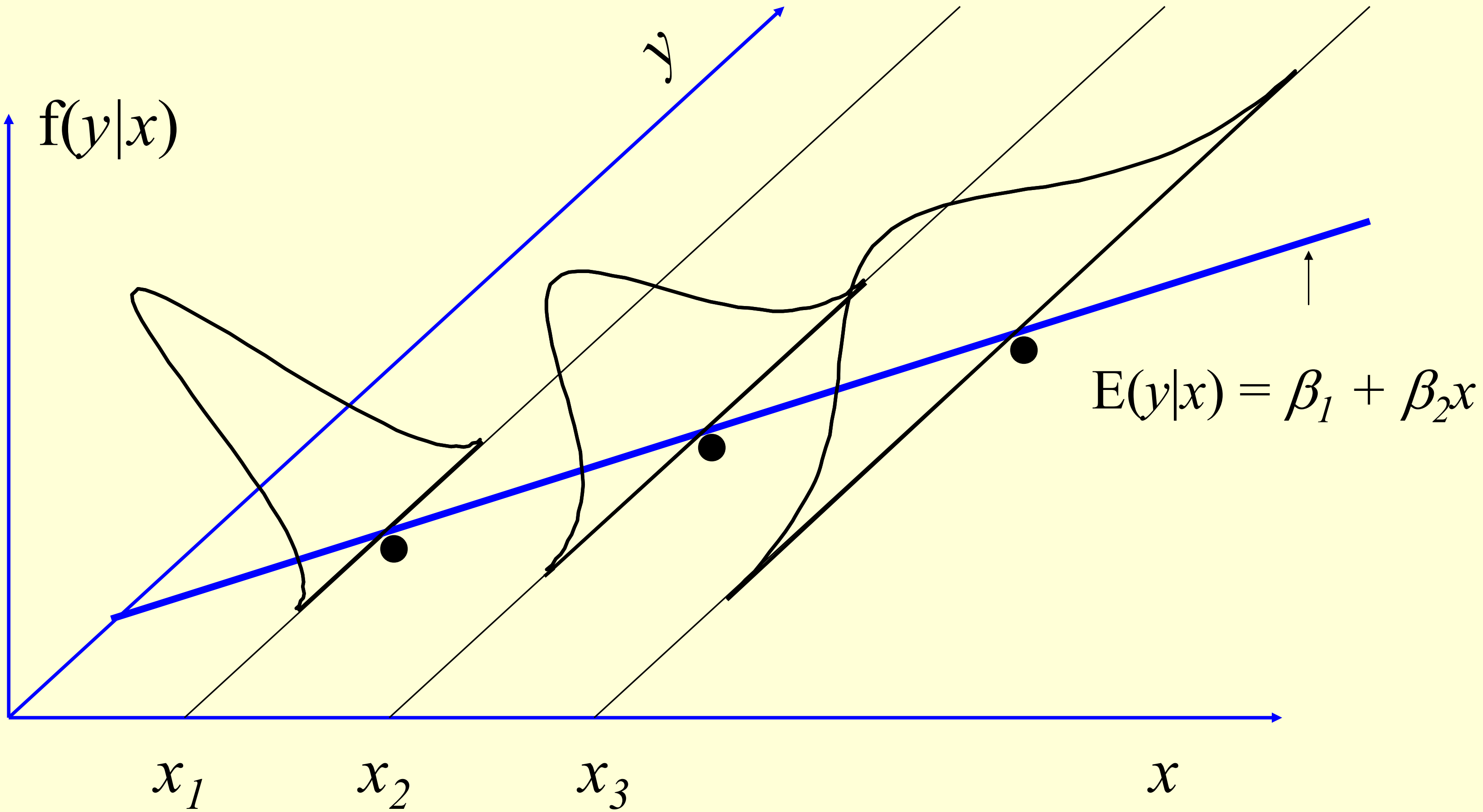
Phương sai bằng nhau



3.2. Mô hình hồi qui đơn

OLS- Hồi qui đơn

Phương sai không bằng nhau



3.2. Mô hình hồi qui đơn

Độ chính xác của SRF

- Từ công thức xác định các tham số của mẫu, ta dễ dàng tính được:

$$Var(\hat{\beta}_2) = \frac{\sigma^2}{\sum (x_i - \bar{x})^2}; \qquad se(\hat{\beta}_2) = \frac{\sigma}{\sqrt{\sum (x_i - \bar{x})^2}}$$

$$Var(\hat{\beta}_1) = \frac{\sum x_i^2}{n\sum (x_i - \bar{x})^2}\sigma^2; \qquad se(\hat{\beta}_1) = \sqrt{\frac{\sum x_i^2}{n\sum (x_i - \bar{x})^2}}\sigma$$

  - Trong đó: $\sigma^2 = Var(ui)$
  - Trong các công thức trên, σ^2 chưa biết, σ^2 được ước lượng bằng ước lượng không chệch của nó là: $\hat{\sigma}^2 = \frac{\sum e_i^2}{n-2}$

Nó chính là độ lệch chuẩn của các giá trị Y quanh đường hồi qui mẫu

Độ phù hợp của mô hình

- Để có thể biết mô hình giải thích được như thế nào hay bao nhiêu % biến động của biến phụ thuộc, người ta sử dụng R^2

$$1 = \frac{\text{ESS}}{TSS} + \frac{\text{RSS}}{TSS} = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2} + \frac{\sum e_i^{\,2}}{\sum(y_i - \bar{y})^2}$$

$$R^2 = \frac{\text{ESS}}{TSS} = 1 - \frac{\text{RSS}}{TSS}$$

Trong đó:

- *TSS là tổng bình phương của tất cả các sai lệch giữa các giá trị quan sát Y_i và giá trị trung bình.*
- *ESS: là tổng bình phương của tất cả các sai lệch giữa các giá trị của biến phụ thuộc Y nhận được từ hàm hồi qui mẫu và giá trị trung bình của chúng. Phần này đo độ chính xác của hàm hồi qui*
- *RSS: là tổng bình phương của tất cả các sai lệch giữa các giá trị quan sát Y và các giá trị nhận được từ hàm hồi qui.*

3.2. Mô hình hồi qui đơn

Độ phù hợp của mô hình

- Tỉ số giữa tổng biến thiên được giải thích bởi mô hình cho tổng bình phương cần được giải thích được gọi là hệ số xác định, hay là trị thống kê “good of fit”. Từ định nghĩa R^2 chúng ta thấy R^2 đo tỷ lệ hay số % của toàn bộ sai lệch Y với giá trị trung bình được giải thích bằng mô hình. Khi đó người ta sử dụng R^2 để đo sự phù hợp của hàm hồi qui; $0 \leq R^2 \leq 1$
  - *R^2 cao nghĩa là mô hình ước lượng được giải thích được một mức độ cao biến động của biến phụ thuộc.*
  - *Nếu R^2 bằng 0. Nghĩa là mô hình không đưa ra thông tin nào về biến phụ thuộc và dự đoán tốt nhất về giá trị của biến phụ thuộc là giá trị trung bình của nó. Các biến "giải thích" thực sự không đưa ra được một giải thích nào.*

3.2. Mô hình hồi qui đơn

Ước lượng khoảng tin cậy của các β_j

- Với các giả thiết đã cho ở phần trước (OLS)- u_i có phân bố $N(0,\sigma^2)$. Nếu thoả mãn thì người ta suy ra:

$$\hat{\beta}_j = N(\beta_j, Var\hat{\beta}_j)$$

$$* \quad T = \frac{\hat{\beta}_j - E(\hat{\beta}_j)}{Se(\hat{\beta}_j)} = \frac{\hat{\beta}_j - \beta_j}{Se(\hat{\beta}_j)} \sim \mathrm{T(n - 2}$$

- Với độ tin cậy $1-\alpha$, ta có ước lượng 2 phía như sau:

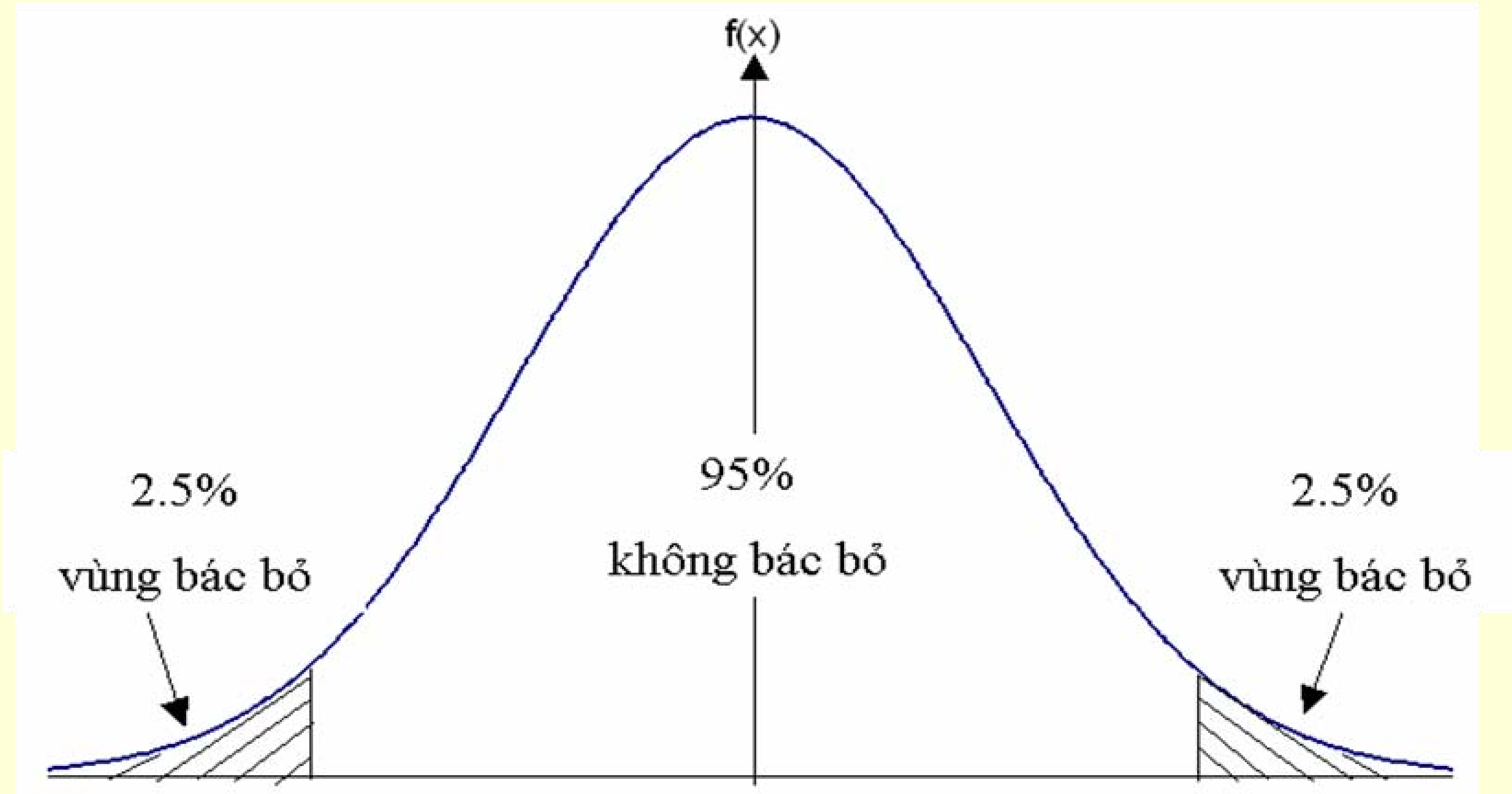
$$P(-t_{\alpha/2}(n-2) \le \frac{\hat{\beta}_j - \beta_j}{Se(\hat{\beta}_j)} \le t_{\alpha/2}(n-2)) = 1-\alpha$$

$$\left[\hat{\beta}_j - t_{\alpha/2}(n-2)Se(\hat{\beta}_j);\ \hat{\beta}_j + t_{\alpha/2}(n-2)Se(\hat{\beta}_j)\right]$$

3.2. Mô hình hồi qui đơn

Ước lượng khoảng tin cậy của các β_j

- Ước lượng 2 phía:



3.2. Mô hình hồi qui đơn

Kiểm định cho các β_j

- Có thể đưa ra giả thiết nào đó đối với β_j, chẳng hạn $\beta_j = \beta_j^*$. Nếu giả thiết này đúng thì:

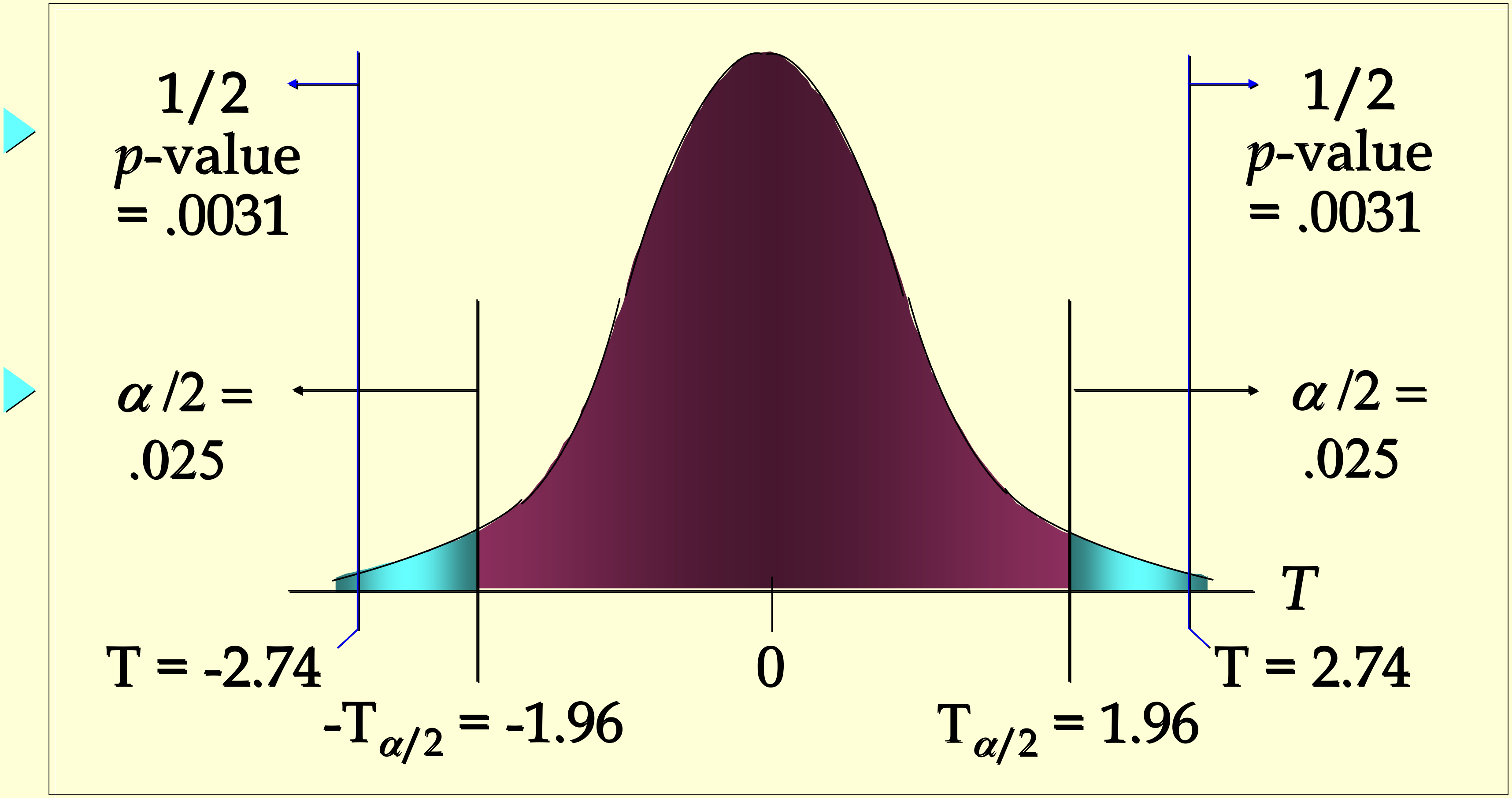
$$T = \frac{\hat{\beta}_j - \beta_j}{Se(\hat{\beta}_j)} \sim T(n - 2)$$

Loại giả thiết	Giả thiết H_0	Giả thiết đối H_1	Miền bác bỏ
Hai phía	$\beta_j = \beta_j^*$	$\beta_j \neq \beta_j^*$	$\lvert t \rvert > t_{\alpha/2}(n-2)$
Phía phải	$\beta_j \leq \beta_j^*$	$\beta_j > \beta_j^*$	$t > t_{\alpha}(n-2)$
Phía trái	$\beta_j \geq \beta_j^*$	$\beta_j < \beta_j^*$	$t < -t_{\alpha}(n-2)$

3.2. Mô hình hồi qui đơn

Kiểm định cho các β_j

- Sử dụng p-value:



3.2. Mô hình hồi qui đơn

Kết quả hồi qui trên Excel

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.9901
R Square	0.9802
Adjusted R Square	0.9798
Standard Error	2.5471
Observations	51

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	15750.3157	15750.3157	2427.7095	0.0000
Residual	49	317.8986	6.4877		
Total	50	16068.2143			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	0.1765	0.4675	0.3775	0.7074	-0.7630	1.1160
income	0.1417	0.0029	49.2718	0.0000	0.1359	0.1474

- *Intercept: Tung độ gốc*
- *Coefficients : Hệ số hồi quy*
- *Standard Error : Sai số chuẩn của ước lượng hệ số*
- *t Stat : Trị thống kê t(n-2)*
- *P-value : Giá trị p*

3.2. Mô hình hồi qui đơn

Kết quả hồi qui trên Eviews

Dependent Variable: WAGE

Method: Least Squares

Included observations: 49

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1120.247	241.4885	4.638923	0.0000
EDUC	112.4522	36.29650	3.098154	0.0033

R-squared	0.169590	Mean dependent var	1820.204
Adjusted R-squared	0.151922	S.D. dependent var	648.2687
S.E. of regression	596.9982	Akaike info criterion	15.66167
Sum squared resid	16751120	Schwarz criterion	15.73888
Log likelihood	-381.7108	F-statistic	9.598561
Durbin-Watson stat	1.582769	Prob(F-statistic)	0.003283

- *C : Tung độ gốc; Coefficient : Hệ số hồi quy*
- *Std. Error : Sai số chuẩn của ước lượng hệ số*
- *t – Statistic : Trị thống kê t(n-k)*
- *Prob: Giá trị p.* *Bác bỏ H_0 khi /t-Statistic/ > $t_{\alpha/2}$ hoặc Prob < α.*

3.2. Mô hình hồi qui đơn

Dự báo

- Dự báo giá trị trung bình của biến phụ thuộc:
  - Giả sử ta biết rằng biến độc lập x và một giá trị x_0 nào đó mà ta cần đưa ra các kết luận về giá trị trung bình của biến phụ thuộc y, thì ta có:

$$E(y|x_0) = E(\beta_1 + \beta_2 x_0 + u_0) = \beta_1 + \beta_2 x_0$$

  - Khi đó đường hồi qui mẫu cho ước lượng điểm $E(y|x_0)$:

$$\hat{y}_0 = \hat{\beta}_1 + \hat{\beta}_2 x_0$$

$\hat{y}_0$ là ước lượng không chệch có phương sai nhỏ nhất của $E(y|x_0)$, tuy nhiên $\hat{y}_0$ vẫn khác giá trị thực của nó.

$\hat{y}_0$ có phân bố chuẩn với kỳ vọng $\beta_1 + \beta_2 x_0$ nên

$$T = \frac{\hat{y}_0 - (\beta_1 + \beta_2 x_0)}{Se(\hat{y}_0)} \sim T(n - 2)$$

3.2. Mô hình hồi qui đơn

Dự báo

- Dự báo giá trị trung bình của biến phụ thuộc:
  - Khoảng tin cậy 1-α của $E(y|x_0)$:

$$P(\hat{\beta}_1 + \hat{\beta}_2 x_0 - t_{\alpha/2}(n-2)Se(\hat{y}_0) \le \beta_1 + \beta_2 x_0$$
$$\le \hat{\beta}_1 + \hat{\beta}_2 x_0 + t_{\alpha/2}(n-2)Se(\hat{y}_0)) = 1-\alpha$$
$$\hat{y}_0 - t_{\alpha/2}(n-2)Se(\hat{y}_0) \le E(y|x_0) \le \hat{y}_0 + t_{\alpha/2}(n-2)Se(\hat{y}_0)$$

  - Trong đó:

$$Var(\hat{y}_0) = \sigma^2\left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum(x_i - \bar{x})^2}\right]$$

3.2. Mô hình hồi qui đơn

Bài tập

- Bảng sau đây cho quan sát theo thời gian về doanh thu bán hàng hàng năm của một công ty (ký hiệu là Y) và chi phí Marketing hàng năm (ký hiệu là X) tính theo giá cố định năm 1990 (đơn vị: tỷ đồng) trong thời kỳ từ 1990-2001.

Y	60.02	86.68	85.66	71.62	88.74	141.27	136.02	132.73	145.48	175.58	158.02	169.81
X	13.44	22.54	18.36	16.8	23.26	40.72	32.75	31.48	37.81	45.29	40.91	46.9

- Từ bảng trên tính được:

$$\sum X_i = 370.26; \quad \sum Y_i = 1451.63; \quad \sum (X_i - \bar{X})^2 = 1500.36$$

$$\sum X_i^2 = 12924.73; \quad \sum (Y_i - \bar{Y})^2 = 17729.63; \quad \sum (X_i - \bar{X})(Y_i - \bar{Y}) = 5077.23$$

$$\sum (Y_i - \hat{Y}_i)^2 = 548.22$$

  - *Ước lượng mô hình Yi =* $\beta_1 + \beta_2 X_i + u_i$
  - *Với hệ số tin cậy là 95% hãy tìm khoảng tin cậy của các hệ số hồi qui.*
  - *Với mức ý nghĩa 5% hãy kiểm định giả thiết* $\beta_2 = 0$*. Từ kết quả nhận được hãy nêu ý nghĩa về mặt kinh tế của kết luận.*
  - *Hãy tính và giải thích ý nghĩa của* R^2
  - *Hãy dự báo doanh thu bán hàng trung bình nếu chi phí Marketing là 50 tỷ đồng.*

3.3. Mô hình hồi qui đa biến (HQ bội)

Giới thiệu mô hình hồi qui đa biến

- Chúng ta đã nghiên cứu mô hình hồi qui đơn. Trong lý thuyết cũng như trong thực tế, có nhiều trường hợp mà biến kinh tế cho không thể giải thích bằng các mô hình hồi qui đơn như vậy.

  Ví dụ:

  - *Lượng cầu phụ thuộc vào giá, thu nhập, giá các hàng hoá khác v.v. Nhớ lại lý thuyết về hành vi người tiêu dùng.*

    QD = f(P, I, Ps, Pc,Market size, T (thị hiếu))
  - *Giá nhà ở phụ thuộc vào diện tích nhà, số phòng ngủ và số phòng tắm ...*
  - *Chi tiêu của hộ gia đình về thực phẩm phụ thuộc vào qui mô hộ gia đình, thu nhập, vị trí địa lý . . .*
  - *Tỷ lệ tử vong trẻ em của quốc gia phụ thuộc vào thu nhập bình quân đầu người, trình độ giáo dục . .*

3.3. Mô hình hồi qui đa biến (HQ bội)

Giới thiệu mô hình hồi qui đa biến

- Khi chúng ta có tập hợp dữ liệu về một biến kinh tế nào đó (biến này được gọi là biến phụ thuộc) và các nhân tố ảnh hưởng đến nó (các nhân tố ảnh hưởng này được gọi là các biến giải thích) thì việc xét đến các ảnh hưởng riêng biệt (hoặc đồng thời) của nhiều nhân tố đến một biến kinh tế có thể được giải thích bằng mô hình hồi qui bội.
- Hàm hồi qui bội tổng thể có dạng

  $y = \beta_1 + \beta_2 x_2 + \beta_3 x_3 + . . . \beta_k x_k + u$ PRF

  Trong đó:

  β_1: là hệ số tự do (hệ số chặn)

  β_j: là hệ số hồi qui riêng

  u: sai số ngẫu nhiên

3.3. Mô hình hồi qui đa biến (HQ bội)

Giả thiết của mô hình hồi qui đa biến

- Các giả thiết OLS cho mô hình hồi qui tuyến tính đơn được giải thích trong mô hình hồi qui bội:

1. $E(u_i) = 0$	Kỳ vọng của các yếu tố ngẫu nhiên u_i bằng 0
2. $Var(u_i) = \sigma^2$	Phương sai bằng nhau với mọi u_i
3. $Cov(u_i, u_j) = 0$	Không có sự tương quan giữa các u_i
4. $Cov(u_i, x_i) = 0$	U và X không tương quan với nhau
5. u_i	Phân phối chuẩn

Giả thiết bổ sung cho mô hình hồi qui bội:

6. Giữa các $x_2, x_3, ..x_k$ không có quan hệ tuyến tính. Nếu $x_2, x_3, ..x_k$ có quan hệ tuyến tính thì người ta nói rằng có hiện tượng đa cộng tuyến.

Hay không tồn tại $\lambda_i \equiv 0$: $\lambda_1 x_{1i} + \lambda_2 x_{2i} + \lambda_3 x_{3i} + ... + \lambda_k x_{ki} = 0$

3.3. Mô hình hồi qui đa biến (HQ bội)

Ước lượng các tham số của mô hình hồi qui đa biến

- Trong thực tế chúng ta thường chỉ có dữ liệu từ mẫu. Từ số liệu mẫu chúng ta ước lượng hồi quy tổng thể.
- Hàm hồi quy mẫu:

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

- Để ước lượng các tham số của mô hình, chúng ta sử dụng phương pháp bình phương nhỏ nhất-OLS (như đã giới thiệu phần trước).
- Các phần dư được định nghĩa giống như trong mô hình hồi qui đơn:

$$e_i = y_i - \hat{y}_i$$

3.3. Mô hình hồi qui đa biến (HQ bội)

Ước lượng các tham số của mô hình hồi qui đa biến

- Chúng ta có thiết lập các điều kiện bậc nhất cho phép tính tối thiểu này như sau:

$$\frac{\partial \sum e_i^2}{\partial \hat{\beta}_1} = -2\sum (y_i - (\hat{\beta}_1 + \hat{\beta}_2 x_{2i} + \hat{\beta}_3 x_{3i} + \ldots + \hat{\beta}_k x_{ki})) = 0$$

$$\frac{\partial \sum e_i^2}{\partial \hat{\beta}_2} = -2\sum (y_i - (\hat{\beta}_1 + \hat{\beta}_2 x_{2i} + \hat{\beta}_3 x_{3i} + \ldots + \hat{\beta}_k x_{ki}))x_{2i} = 0$$

....

$$\frac{\partial \sum e_i^2}{\partial \hat{\beta}_k} = -2\sum (y_i - (\hat{\beta}_1 + \hat{\beta}_2 x_{2i} + \hat{\beta}_3 x_{3i} + \ldots + \hat{\beta}_k x_{ki}))x_{ki} = 0$$

- *Hệ phương trình mà chúng ta có được gọi là hệ phương trình chuẩn. Chúng ta có thể giải k phương trình chuẩn này để tìm k hệ số beta chưa biết.*
- *Sự trình bày đơn giản nhất của lời giải này ở dưới dạng đại số ma trận. Tuy nhiên sử dụng phần mềm EViews hay các phần mềm phân tích dữ liệu khác chúng ta có thể tìm dễ dàng các hệ số hồi qui.*

3.3. Mô hình hồi qui đa biến (HQ bội)

Hệ số xác định bội R^2 và hệ số xác định bội đã hiệu chỉnh

- Trong mô hình hồi qui hai biến R^2 đo độ thích hợp của hàm hồi qui. Nó chính là tỷ lệ của toàn bộ sự biến đổi của biến phụ thuộc y do biến giải thích x gây ra. Trong mô hình hồi qui bội tỷ lệ của toàn bộ sự khác biệt của biến y do tất cả các biến X gây ra được gọi là hệ số xác định bội, ký hiệu là R^2:

$$R^2 = 1 - \frac{\sum e_i^2}{\sum (y_i - \bar{y})^2}$$

- $0 \leq R^2 \leq 1$. Nếu $R^2 = 1$, có nghĩa là đường hồi qui giải thích 100% thay đổi của y. Nếu $R^2 = 0$, có nghĩa là mô hình không giải thích sự thay đổi nào của y.
- R^2 Là hàm không giảm của số biến giải thích có trong mô hình, do đó nếu tăng số biến giải thích có trong mô hình thì R^2 cũng tăng. Vấn đề cần đặt ra là khi nào cần đưa thêm biến giải thích mới vào trong mô hình?

3.3. Mô hình hồi qui đa biến (HQ bội)

Hệ số xác định bội R^2 và hệ số xác định bội đã hiệu chỉnh

- Để ngăn chặn tình trạng "có đưa thêm biến vào mô hình" như đã nêu trên, một phép đo khác về mức độ thích hợp được sử dụng thường xuyên hơn. Phép đo này gọi là **R^2 hiệu chỉnh** hoặc **R^2 hiệu chỉnh theo bậc tự do** (kết quả này luôn được in ra khi thực hiện hồi qui bằng những phần mềm chuyên dụng). Để phát triển phép đo này, trước hết phải nhớ là R^2 đo lường tỷ số giữa phương sai của Y "được giải thích" bằng mô hình; một cách tương đương, nó bằng 1 trừ đi tỷ số "không được giải thích" do phương sai của sai số Var(*u*). Ta có thể biểu diễn công thức tính như sau:

$$\overline{R}^2 = 1 - \frac{Var(u)}{Var(Y)}$$

3.3. Mô hình hồi qui đa biến (HQ bội)

Hệ số xác định bội R^2 và hệ số xác định bội đã hiệu chỉnh

- Chúng ta biết rằng một ước lượng không chệch của Var (u) được tính bằng RSS/(n – k), và một ước lượng không chệch của Var (Y) được tính bằng TSS/(n – 1). Thay vào phương trình trên ta có:

$$\overline{R}^2 = 1 - \frac{RSS/(n-k)}{TSS/(n-1)} = 1 - \frac{RSS(n-1)}{TSS(n-k)}$$

$$= 1 - \frac{(n-1)}{(n-k)}(1-R^2) = 1 - \frac{\hat{\sigma}^2(n-1)}{TSS}$$

- Qua thao tác hiệu chỉnh này thì chỉ những biến thực sự làm tăng khả năng giải thích của mô hình mới xứng đáng được đưa vào mô hình .

3.3. Mô hình hồi qui đa biến (HQ bội)

Ước lượng khoảng tin cậy và kiểm định cho β_j

- ***Ước lượng 2 phía***, ta tìm được $t_{\alpha/2}$ (n-k) thoả mãn.

$$P(-t_{\alpha/2}(n-k) \le \frac{\hat{\beta}_j - \beta_j}{Se(\hat{\beta}_j)} \le t_{\alpha/2}(n-k)) = 1-\alpha$$

- Do vậy khoảng tin cậy là:

$$\left[\hat{\beta}_j - t_{\alpha/2}(n-k)Se(\hat{\beta}_j);\ \hat{\beta}_j + t_{\alpha/2}(n-k)Se(\hat{\beta}_j)\right]$$

- Chúng ta cũng có thể kiểm định giả thiết $\beta_j = \beta_j^*$, Thực hiện tương tự như hồi qui đơn.

3.3. Mô hình hồi qui đa biến (HQ bội)

Kiểm định ý nghĩa của hàm hồi qui

- Trong mô hình hồi qui đa biến, giả thiết "không" cho rằng mô hình không có ý nghĩa được hiểu là tất cả các hệ số hồi qui riêng (các tham số độ dốc) đều bằng không. Ứng dụng kiểm định Wald (thường được gọi là kiểm định *F*) được tiến hành cụ thể như sau:
  - **Bước 1** *Giả thuyết không* là H_0: $\beta_2 = \beta_3 = \ldots = \beta_k = 0$

    Giả thuyết ngược lại là H_1: có ít nhất một trong những giá trị β không bằng không.
  - **Bước 2** Trước tiên hồi qui Y theo một số hạng không đổi và $X_2, X_3, .., X_k$, sau đó tính tổng bình phương sai số RSS_U. Kế đến tính RSS_R. Chúng ta đã định nghĩa phân phối *F* là tỷ số của hai biến ngẫu nhiên phân phối chi bình phương độc lập. Điều này cho ta trị thống kê:

$$F_c = \frac{\left[RSS_R - RSS_U\right]/(k-m)}{RSS_U/(n-k)} \sim F(\alpha, k-m, n-k)$$

3.3. Mô hình hồi qui đa biến (HQ bội)

Kiểm định ý nghĩa của hàm hồi qui

- Vì H_0: $\beta_2 = \beta_3 = \ldots = \beta_k = 0$, dễ dàng thấy rằng trị thống kê kiểm định đối với giả thiết này sẽ là:

$$F_c = \frac{\left[RSS_R - RSS_U\right]/(k-m)}{RSS_U\,/(n-k)} = \frac{ESS\,/(k-1)}{RSS\,/(n-k)} \sim F(\alpha, k-m, n-k)$$

- **Bước 3** Từ số liệu trong bảng *F* tương ứng với bậc tự do *k − 1* cho tử số và *n − k* cho mẫu số, và với mức ý nghĩa cho trước α, ta có *F*(α, k-1,n-k)* sao cho diện tích bên phải của *F** là α.
- **Bước 4** Bác bỏ *giả thuyết không* ở mức ý nghĩa a nếu *Fc > F**. Đối với phương pháp giá trị *p*, tính giá trị $p = P(F > Fc|H_0)$ và bác bỏ *giả thuyết không* nếu giá trị *p nhỏ hơn* mức ý nghĩa.

3.3. Mô hình hồi qui đa biến (HQ bội)

Bài tập

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.5676
R Square	0.3222
Adjusted R Square	0.2770
Standard Error	551.2095
Observations	49

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	6499678.255	2166559	7.130784	0.000510121
Residual	45	13672433.7	303831.9		
Total	48	20172111.96			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	632.244	423.379	1.493	0.142	-220.484	1484.972
EDUC	142.510	34.859	4.088	0.000	72.299	212.720
EXPER	43.225	14.304	3.022	0.004	14.417	72.034
AGE	-1.913	8.394	-0.228	0.821	-18.819	14.992

- *Hãy giải thích dấu mà anh/chị mong muốn cho các hệ số β_2 , β_3, β_4*
- *Giải thích ý nghĩa của các hệ số hồi quy.*
- *Hãy cho biết mô hình nhận hồi qui có ý nghĩa với αbằng 5% hay không*

Chương 4

PHƯƠNG PHÁP DỰ BÁO ĐỊNH LƯỢNG

4.1. Giới thiệu

Khái niệm và vai trò của dự báo

- Dự báo theo tiếng Hy Lạp là Prognosis - sự tiên đoán, sự thấy trước.
- Dự báo là sự tiên đoán có căn cứ khoa học, mang tính chất xác suất về mức độ, nội dung, các mối quan hệ, trạng thái, xu hướng phát triển của đối tượng nghiên cứu hoặc về cách thức và thời hạn đạt được các mục tiêu nhất định đã đề ra trong tương lai.
- *Tiên đoán khoa học:* đây là tiên đoán dựa trên việc phân tích mối quan hệ qua lại giữa các đối tượng trong khuôn khổ của một hệ thống lý luận khoa học nhất định. Nó dựa trên việc phân tích tính quy luật phát triển của đối tượng dự báo và các điều kiện ban đầu với tư cách như là các giả thiết. Tiên đoán khoa học là kết quả của sự kết hợp giữa những phân tích định tính và những phân tích định lượng các quá trình cần dự báo.

4.1. Giới thiệu

Khái niệm và vai trò của dự báo

- Dự báo là một yếu tố quan trọng của hầu hết các quyết định kinh doanh và lập kế hoạch kinh tế.
- Công tác dự báo là vô cùng quan trọng bởi lẽ nó cung cấp các thông tin cần thiết nhằm phát hiện và bố trí sử dụng các nguồn lực trong tương lai một cách có căn cứ thực tế. Với những thông tin mà dự báo đưa ra cho phép các nhà hoạch định chính sách có những quyết định về đầu tư, các quyết định về sản xuất, về tiết kiệm và tiêu dùng, các chính sách tài chính, chính sách kinh tế vĩ vô.
- Hầu như mọi lĩnh vực chức năng của doanh nghiệp đều sử dụng một loại dự báo nào đó, ví dụ:

4.1. Giới thiệu

Khái niệm và vai trò của dự báo

- Kế toán: dự báo chi phí và doanh thu trong kế hoạch nộp thuế.
- Phòng nhân sự: dự báo nhu cầu tuyển dụng và những thay đổi trong công sở.
- Chuyên gia tài chính: dự báo ngân lưu.
- Quản đốc sản xuất: dự báo nhu cầu nguyên vật liệu và tồn kho.
- Giám đốc marketing: Dự báo doanh số để thiết lập ngân sách cho quảng cáo.

4.1. Giới thiệu

Qui trình dự báo

- Bước 1. Xác định mục đích
- Bước 2. Xác định khoảng thời gian dự báo
- Bước 3. Chọn phương pháp dự báo
- Bước 4. Thu thập và phân tích dữ liệu
- Bước 5. Tiến hành dự báo
- Bước 6. Kiểm chứng kết quả và rút kinh nghiệm

4.1. Giới thiệu

Các phương pháp dự báo

- **PHƯƠNG PHÁP DỰ BÁO**
  - **PHƯƠNG PHÁP ĐỊNH TÍNH**
    - -Lấy ý kiến của ban lãnh đạo
    - -Lấy ý kiến của bộ phận bán hàng
    - -Phương pháp lấy ý kiến của người tiêu dùng
    - -Phương pháp chuyên gia
  - **PHƯƠNG PHÁP ĐỊNH LƯỢNG**
    - **Các mô hình nhân quả**
      - -Hồi quy
      - -Phân tích tương quan
    - **Các mô hình chuỗi thời gian**
      - -Bình quân đơn giản
      - -Bình quân di động
      - -San bằng số mũ
      - -Chuỗi thời gian
      - -Phương pháp Box- Jenkins

Các phương pháp dự báo

4.1. Giới thiệu

Dự báo định lượng

- Các phương pháp dự báo định lượng dựa vào các số liệu thống kê và thông qua các công thức toán học được thiết lập để dự báo cho tương lai. Khi dự báo, nếu không xét đến các nhân tố ảnh hưởng khác có thể dùng các phương pháp dự báo theo dãy số thời gian. Nếu cần ảnh hưởng của các nhân tố khác đến nhu cầu có thể dùng các mô hình nhân quả (hồi quy, tương quan).
- Ưu điểm của dự báo định lượng:
    - Kết quả dự báo hoàn toàn khách quan
    - Có phương pháp đo lường độ chính xác dự báo
    - Ít tốn thời gian để tìm ra kết quả dự báo
    - Có thể dự báo điểm hay dự báo khoảng

4.1. Giới thiệu

Nguồn dữ liệu

- Tùy vào phương pháp dự báo được chọn:
  - Một số phương pháp chỉ cần chuỗi số liệu sẽ được dự báo: như dự báo thô, phân tích, san mũ, ARIMA
  - Các phương pháp hồi qui yêu cầu phải có số liệu cho mỗi biến sử dụng trong mô hình
- Số liệu nội bộ của tổ chức
- Số liệu bên ngoài tổ chức

4.2. Dự báo dựa trên dữ liệu chuỗi thời gian

Khái quát

- Các hiện tượng kinh tế - xã hội luôn luôn biến động qua thời gian. Để nghiên cứu sự biến động này người ta dùng phương pháp chuỗi thời gian (dãy số thời gian). *Chuỗi thời gian là dãy các trị số của một chỉ tiêu nào đó được sắp xếp theo thứ tự thời gian.*
- Phương pháp dự báo theo dãy số thời gian được xây dựng trên một giả thiết về sự tồn tại và lưu lại các nhân tố quyết định đại lượng dự báo từ quá khứ đến tương lai. Trong phương pháp này đại lượng cần dự báo được xác định trên cơ sở phân tích chuỗi các số liệu thống kê được trong quá khứ .
- (vd.: số liệu về nhu cầu sản phẩm, doanh thu, lợi nhuận, chi phí, năng suất hay chỉ số tiêu dùng…).

4.2. Dự báo dựa trên dữ liệu chuỗi thời gian

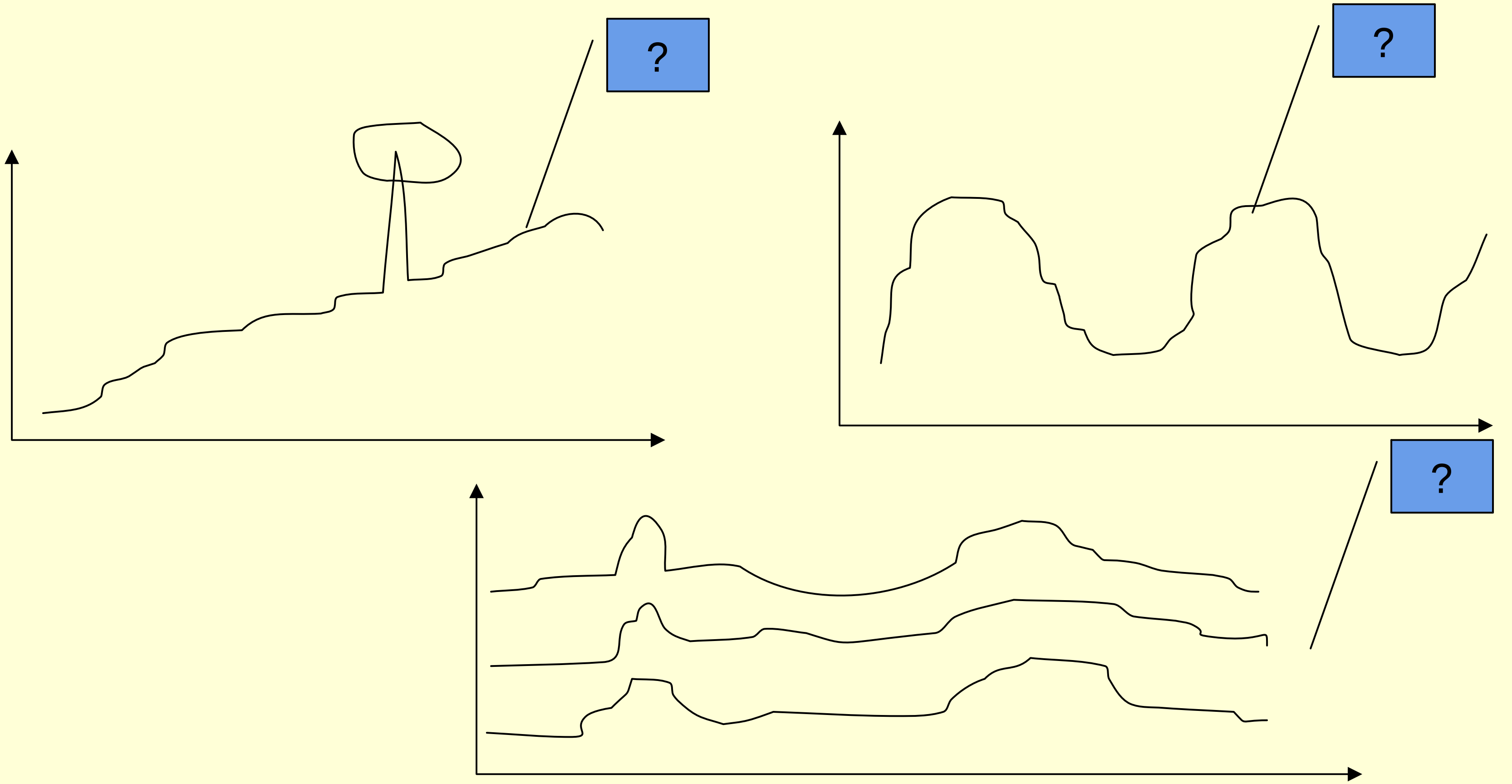
Khái quát

- Một số tính chất của chuỗi thời gian
  - Tính xu hướng (trend);
  - Tính thời vụ (seasonality);
  - Tính chu kỳ (cycles);
  - Những biến động ngẫu nhiên (random variation).

4.2. Dự báo dựa trên dữ liệu chuỗi thời gian

Khái quát

- Một số tính chất của chuỗi thời gian



4.2. Dự báo dựa trên dữ liệu chuỗi thời gian

Phương pháp dự báo giản đơn

- Phương pháp trung bình giản đơn là phương pháp dự báo trên cơ sở lấy trung bình của các dữ liệu đã qua, trong đó các giá trị của các giai đoạn trước đều có trọng số như nhau.
- Nội dung:
  - Dự báo giá trị ở kỳ tiếp theo (t) sẽ bằng chính giá trị của kỳ trước đó (t-1).
- Công thức: $F_t = D_{t-1}$ (4-1)
  - *Trong đó:*
  - *F_t - mức dự báo ở kỳ t;*
  - *D_{t-1} – giá trị thực tế của kỳ t-1*
- Ưu điểm: Đơn giản, có thể ứng dụng hiệu quả trong trường hợp chuỗi có xu hướng rõ ràng.
- Nhược điểm: Mức độ chính xác của dự báo thấp.

4.2. Dự báo dựa trên dữ liệu chuỗi thời gian

Phương pháp trung bình giản đơn

- Công thức:

$$F_t = \frac{\sum_{i=1}^{t-1} D_i}{n}, \qquad (4\text{-}2)$$

- Trong đó:
  - F_t – là giá trị dự báo cho giai đoạn t;
  - D_i _ là giá trị thực tế của giai đoạn i;
  - n – số giai đoạn thực tế dùng để quan sát (n=t-1).
- Ưu điểm:
  - Chính xác hơn phương pháp giản đơn
  - Phù hợp với những dòng yêu cầu đều có xu hướng ổn định.
- Nhược điểm:
  - Phải lưu trữ một số lượng dữ liệu khá lớn.

4.2. Dự báo dựa trên dữ liệu chuỗi thời gian

Phương pháp trung bình giản đơn

- Ví dụ 1: Hãy dự báo nhu cầu tháng tới dựa trên mức bán hàng thực tế của các tháng trước:

Tháng	**Mức bán thực tế (Dt)**	**Dự báo (Ft)**
1	100	--
2	110	F2=D2=100
3	120	F3=(D1+D2)/2=105
4	115	F4=110
5		F5=?

4.2. Dự báo dựa trên dữ liệu chuỗi thời gian

Phương pháp trung bình động (TB trượt)

- Trong trường hợp khi có sự biến động, trong đó thời gian gần nhất có ảnh hưởng nhiều nhất đến kết quả dự báo, thời gian càng xa thì ảnh hưởng càng nhỏ ta dùng phương pháp trung bình động sẽ thích hợp hơn.
- Dự báo cho giai đoạn tiếp theo dựa trên cơ sở kết quả trung bình của các kỳ trước đó thay đổi (trượt) trong một giới hạn thời gian nhất định.
- Công thức:

$$F_t = \frac{\sum_{i=1}^{n} D_{t-i}}{n} \qquad (4\text{-}3)$$

- Trong đó:
  - F_t – là giá trị dự báo cho giai đoạn t;
  - D_{t-i} – là giá trị thực tế của giai đoạn t-i;
  - n – số giai đoạn quan sát.

4.2. Dự báo dựa trên dữ liệu chuỗi thời gian

Phương pháp trung bình động (TB trượt)

- Ưu điểm: Cho độ chính xác tương đối, Rút ngắn số liệu lưu trữ
- Nhược điểm: Không cho thấy được mối tương quan trong các đại lượng của dòng yêu cầu.
- Ví dụ 2: Dự báo nhu cầu cho các tháng tới bằng phương pháp trung bình động, với n=3.

Tháng	Mức bán thực tế (Dt)	Dự báo (Ft)
1	100	
2	110	
3	120	
4	115	F4=(120 + 110 +100)/3
5	125	F5=(115 + 120 + 110)/3
6		F6=?

4.2. Dự báo dựa trên dữ liệu chuỗi thời gian

Phương pháp trung bình động có trọng số

- Là phương pháp trung bình động có tính đến ảnh hưởng của từng giai đoạn khác nhau đến biến dự báo thông qua sử dụng trọng số
- Công thức:

$$F_t = \sum_{i=1}^{n} D_{t-i} \cdot \alpha_{t-i} \qquad (4\text{-}4)$$

- Trong đó:
  - Dt-i – là giá trị thực ở giai đoạn t-i
  - α_{t-i} – là trọng số của giai đoạn t-i với $\sum \alpha_{t-i} = 1$ và $0 \leq \alpha_{t-i} \leq 1$.
- Ưu điểm: Cho kết quả sát với thực tế hơn so với pp tbd giản đơn vì có sử dụng hệ số.
- Nhược điểm:
  - *Dự báo không bắt kịp xu hướng thay đổi của biến;*
  - *Đòi hỏi ghi chép số liệu chính xác và đủ lớn.*

4.2. Dự báo dựa trên dữ liệu chuỗi thời gian

Phương pháp san bằng hàm số mũ

- Nhằm khắc phục nhược điểm của phương pháp trước, pp san bằng mũ cho rằng dự báo mới bằng dự báo của giai đoạn trước đó cộng với tỉ lệ chênh lệch giữa giá trị thực và dự báo của giai đoạn đó qua, có điều chỉnh cho phù hợp.
- Công thức:

$$F_t = F_{t-1} + \alpha\left(D_{t-1} - F_{t-1}\right) = \alpha D_{t-1} + \left(1-\alpha\right)F_{t-1} \qquad (4\text{-}5)$$

- Trong đó:
  - *F_t – Dự báo nhu cầu giai đoạn t*
  - *F_{t-1} - Dự báo nhu cầu giai đoạn t-1*
  - *D_{t-1} – Nhu cầu thực của giai đoạn t-1*
  - *α- Hệ số san bằng mũ*

4.2. Dự báo dựa trên dữ liệu chuỗi thời gian

Phương pháp san bằng hàm số mũ

- Vì sao lại gọi là pp san bằng hàm số mũ?. Ta thấy rằng:

$$F_t = \alpha D_{t-1} + (1-\alpha)F_{t-1}$$

$$\Leftrightarrow F_t = \alpha D_{t-1} + (1-\alpha)\left[\alpha D_{t-2} + (1-\alpha)F_{t-2}\right]$$

$$\Leftrightarrow F_t = \alpha D_{t-1} + \alpha(1-\alpha)D_{t-2} + \alpha(1-\alpha)^2 D_{t-3} + \alpha(1-\alpha)^3 D_{t-4} + \cdots$$

Nhận xét:

- *Ảnh hưởng của các số liệu trong quá khứ đối với kết quả dự báo có giá trị giảm dần với một trọng số như nhau là (1-α) -> α - được gọi là hệ số san bằng hàm số mũ.*

4.2. Dự báo dựa trên dữ liệu chuỗi thời gian

Phương pháp san bằng hàm số mũ

- Chọn α như thế nào?:
  - Chỉ số α thể hiện độ nhảy cảm của sai số dự báo, nên phụ thuộc nhiều vào loại hình sản phẩm và kinh nghiệm của người khảo sát;
  - $0 \leq \alpha \leq 1$, người ta thường chọn α [0.05-0.5];
  - Để có α phù hợp phải dùng phương pháp thử nghiệm và chọn kết quả có sai số nhỏ nhất.

4.2. Dự báo dựa trên dữ liệu chuỗi thời gian

Phương pháp san bằng hàm số mũ

- Ví dụ: Dự báo với số liệu trong Ví dụ 2

Tháng i	Nhu cầu thực tế (Dt)	Nhu cầu dự báo (Ft)			
		$\alpha = 0.10$		$\alpha = 0.40$	
		$F_{t,0.1}$	Sai số	$F_{t,0.4}$	Sai số
1	100	-	-	-	-
2	110				
3	120				
4	115				
5	125				
6					

4.2. Dự báo dựa trên dữ liệu chuỗi thời gian

Phương pháp san bằng hàm số mũ

- Ví dụ: Dự báo với số liệu trong Ví dụ 2

Tháng i	Nhu cầu thực tế (Dt)	Nhu cầu dự báo (Ft)			
		$\alpha = 0.10$		$\alpha = 0.40$	
		$F_{t,0.1}$	Sai số	$F_{t,0.4}$	Sai số
1	100	-	-	-	-
2	110	100	10	100	10
3	120	101	19	104	16
4	115	102.9	12.1	110.4	4.6
5	125	104.11	20.89	112.24	12.76
6		106.20		117.34	

4.3. Dự báo bằng phương pháp dự báo nhân quả

Khái niệm

- Là phương pháp dự báo dựa trên việc xác định mối quan hệ giữa các đại lượng (biến), rồi dựa vào đó để đưa ra dự báo.
- Ví dụ: Doanh thu & chi phí; quảng cáo & lợi nhuận; giá cả & tiền lương.
- Ta sẽ tìm hiểu hai phương pháp cơ bản: *hồi qui tuyến tính* và *phân tích tương quan*.

4.3. Dự báo bằng phương pháp dự báo nhân quả

Phân tích tương quan

- Phân tích tương quan đánh giá mối quan hệ giữa 2 nhân tố. Giá trị cuối cùng (hệ số tương quan) chỉ ra liệu có sự thay đổi của nhân tố này sẽ dẫn đến thay đổi trong nhân tố kia hay không.
- Một hệ số tương quan thấp (ví dụ: <0.10) chỉ ra rằng mối quan hệ giữa 2 nhân tố rất yếu hoặc không tồn tại. Một hệ số tương quan cao (gần bằng ±1) chỉ ra rằng nhân tố phụ thuộc sẽ thay đổi khi nhân tố độc lập thay đổi.
- *Ví dụ: Phân tích tương quan khá hữu ích trong việc đánh giá mối quan hệ giữa 2 loại chứng khoán. Thông thường, một mức giá an toàn dẫn dắt hoặc dự báo giá cả của một loại chứng khoán khác. Hay, hệ số tương quan của vàng đối nghịch với đồng đô la. Điều này có nghĩa là một sự gia tăng trong đồng đô la thường báo trước một sự giảm giá của vàng.*

4.3. Dự báo bằng phương pháp dự báo nhân quả

Phân tích tương quan

- Nếu có số liệu về hai đại lượng x, y. Để đánh giá mức độ quan hệ giữa hai đại lượng này, người ta sử dụng *hệ số tương quan ρ*, được tính như sau:

$$\rho_{xy} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}} \frac{\mathrm{cov}(X, Y)}{\sqrt{\mathrm{var}(X)\,\mathrm{var}(Y)}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

- Hoặc chúng ta cũng có thể dễ dàng tính được thông qua công cụ Correlation trong Excel

4.3. Dự báo bằng phương pháp dự báo nhân quả

Dự báo bằng mô hình hồi qui

- Biểu diễn mối quan hệ giữa hai đại lượng thông qua phương trình hồi qui:

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

- Chúng ta sử dụng phương pháp bình phương nhỏ nhất-OLS để ước lượng các tham số của mô hình.
- Dựa vào phương trình hồi qui để đưa ra dự báo.
- Nhận xét, đánh giá về kết quả dự báo

4.3. Dự báo bằng phương pháp dự báo nhân quả

Dự báo bằng mô hình hồi qui

- Dự báo giá trị trung bình của biến phụ thuộc:
  - Giả sử ta biết rằng biến độc lập x và một giá trị x_0 nào đó mà ta cần đưa ra các kết luận về giá trị trung bình của biến phụ thuộc y, thì ta có:

$$E(y|x_0) = E(\beta_1 + \beta_2 x_0 + u_0) = \beta_1 + \beta_2 x_0$$

  - Khi đó đường hồi qui mẫu cho ước lượng điểm $E(y|x_0)$:

$$\hat{y}_0 = \hat{\beta}_1 + \hat{\beta}_2 x_0$$

$\hat{y}_0$ là ước lượng không chệch có phương sai nhỏ nhất của $E(y|x_0)$, tuy nhiên $\hat{y}_0$ vẫn khác giá trị thực của nó.

$\hat{y}_0$ có phân bố chuẩn với kỳ vọng $\beta_1 + \beta_2 x_0$ nên

$$T = \frac{\hat{y}_0 - (\beta_1 + \beta_2 x_0)}{Se(\hat{y}_0)} \sim T(n-2)$$

4.3. Dự báo bằng phương pháp dự báo nhân quả

Dự báo bằng mô hình hồi qui

- Dự báo giá trị trung bình của biến phụ thuộc:
  - Khoảng tin cậy 1-α của $E(y|x_0)$:

$$P(\hat{\beta}_1 + \hat{\beta}_2 x_0 - t_{\alpha/2}(n-2)Se(\hat{y}_0) \leq \beta_1 + \beta_2 x_0$$

$$\leq \hat{\beta}_1 + \hat{\beta}_2 x_0 + t_{\alpha/2}(n-2)Se(\hat{y}_0)) = 1 - \alpha$$

$$\hat{y}_0 - t_{\alpha/2}(n-2)Se(\hat{y}_0) \leq E(y|x_0) \leq \hat{y}_0 + t_{\alpha/2}(n-2)Se(\hat{y}_0)$$

  - Trong đó:

$$Var(\hat{y}_0) = \sigma^2 \left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right]$$

Chương 5

MÔ HÌNH TOÁN KINH TẾ VÀ PHƯƠNG PHÁP TỐI ƯU

5.1. Tổng quan

Khái niệm về mô hình kinh tế và mô hình toán kinh tế

- Mô hình kinh tế:
  - Mô hình kinh tế là mô hình phản ánh các đối tượng trong lĩnh vực hoạt động kinh tế.
  - Để xây dựng mô hình kinh tế, chúng ta còn cần phải thu thập, xử lý các thông tin về những kết quả nghiên cứu liên quan, các dữ liệu đã được công bố và các kiến thức của các ngành khoa học khác.
  - Người ta thường mô tả và phân tích các hiện tượng, hệ thống kinh tế – xã hội dưới ba nhóm mô hình kinh tế sau:
    - *Mô hình kinh tế vi mô (Micro).*
    - *Mô hình kinh tế vĩ mô (Macro).*
    - *Mô hình kinh tế phát triển.*
  - Ngày nay, các mô hình kinh tế được diễn đạt và phân tích bằng ngôn ngữ, tư duy và công cụ toán học, một khoa học chặt chẽ, chính xác có khả năng diễn tả và phân tích một cách đầy đủ, bản chất, khái quát nhất sự vận động và phát triển của các hiện tượng, hệ thống kinh tế - xã hội.

5.1. Tổng quan

Khái niệm về mô hình kinh tế và mô hình toán kinh tế

- Mô hình toán kinh tế:
  - Mô hình toán kinh tế là mô hình kinh tế được diễn đạt bằng ngôn ngữ toán học.
  - Bản chất của quá trình mô hình hoá một hiện tượng, một hệ thống kinh tế là mô hình hoá quá trình vận động của nó, nghĩa là xây dựng phương trình trạng thái cho nó. Để xây dựng mô hình toán học của một hiện tượng, một hệ thống kinh tế cụ thể, ta phải chọn các biến kinh tế cho nó, đó là các biến điều khiển, các biến ngẫu nhiên (gọi chung là các biến vào) và các biến trạng thái, các biến ra (kết quả sản xuất), sau đó mô tả quan hệ giữa các biến đó bằng những hệ thức toán học.

5.1. Tổng quan

Cấu trúc của mô hình toán kinh tế

- Cấu trúc của mô hình toán kinh tế thường gồm hai bộ phận chính: các biến và các ràng buộc nhằm diễn tả chặt chẽ, chính xác, đầy đủ hơn các hiện tượng và hệ thống kinh tế đang nghiên cứu, người ta đưa thêm vào mô hình phần giả thiết và chú thích hoặc nhận xét.
  a) *Các biến kinh tế của mô hình:* Các biến kinh tế là các đại lượng biến thiên đặc trưng cho các yếu tố cơ bản của các hiện tượng kinh tế và hệ thống kinh tế ta cần nghiên cứu. Các biến kinh tế thay đổi giá trị trong phạm vi nhất định. Tuỳ theo mục đích nghiên cứu cũng như khả năng về nguồn dữ liệu liên quan, biến kinh tế trong mô hình toán kinh tế thường được phân ra làm ba loại
    - *Các biến ngoại sinh (Biến giải thích, biến độc lập).*
    - *Các biến nội sinh (biến được giải thích, biến phụ thuộc).*
    - *Các tham số (thông số)*

5.1. Tổng quan

Cấu trúc của mô hình toán kinh tế

b) *Các ràng buộc của mô hình :*

- Các ràng buộc của mô hình là các hệ thức toán học phản ánh mối quan hệ kinh tế, quan hệ hành vi, quan hệ cơ học, quan hệ kỹ thuật, quan hệ đồng nhất,... giữa các yếu tố kinh tế trong hệ thống kinh tế hoặc hiện tượng kinh tế mà ta đồng nghiên cứu.
- Hình thức biểu hiện mối quan hệ giữa các biến kinh tế là các phương trình, bất phương trình. Ngoài ra các quan hệ kinh tế còn được biểu hiện bởi các hệ thức toán học khác. Các phương trình trong mô hình gọi là phương trình cấu trúc. Phương trình cấu trúc ở dạng đơn giản là những hàm số (hàm sản xuất, hàm kinh tế), ở dạng phức tạp hơn là những phương trình, hệ phương trình đại số, vi phân hoặc sai phân

5.1. Tổng quan

Xây dựng mô hình toán kinh tế

- Việc mô hình hoá toán học các hiện tượng hoặc một hệ thống kinh tế thường được tiến hành theo bốn bước sau.
  - Bước 1: Xây dựng mô hình định tính cho đối tượng kinh tế cần nghiên cứu, nghĩa là xác định các yếu tố có ý nghĩa quan trọng nhất và xác lập các qui luật mà các yếu tố kinh tế phải tuân theo. Nói cách khác là phát biểu mô hình bằng lời, bằng biểu đồ cùng các điều kiện kinh tế, kỹ thuật, xã hội, tự nhiên và các mục tiêu cần đạt được.
  - Bước 2: Xây dựng mô hình toán học cho đối tượng kinh tế cần nghiên cứu, nghĩa là diễn tả lại dưới dạng ngôn ngữ toán học cho mô hình định tính, bao gồm xác định biến kinh tế và các ràng buộc của các biến kinh tế.

5.1. Tổng quan

Xây dựng mô hình toán kinh tế

- Bước 3; Sử dụng các công cụ toán học để khảo sát và giải quyết mô hình toán học đã xác lập ở bước 2. Căn cứ vào mô hình đã xây dựng, lựa chọn hoặc xây dựng phương pháp giải cho phù hợp. Tiếp đó cụ thể hoá phương pháp bằng các thuật toán tối ưu và thể nghiệm giải bài toán trên máy tính điện tử.
- Bước 4: Dựa vào các số liệu thu thập được, mô phỏng trên máy tính các tình huống trong quá khứ và hiện tại, dự đoán và kiểm định sự phù hợp của mô hình đối với lý luận và thực tiễn.

5.1. Tổng quan

Sử dụng mô hình toán kinh tế trong nghiên cứu lựa chọn giải pháp tối ưu

Sau khi đã xây dựng và hiệu chỉnh mô hình phù hợp với hiện tượng và quá trình kinh tế, ta có thể sử dụng mô hình để phân tích động thái và hành vi của đối tượng kinh tế từ đó lựa chọn giải pháp tốt nhất cho quá trình quản lý điều khiển kinh tế.

- ***Sử dụng mô hình kinh tế Vi mô (Micro)***.
    - Người ta sử dụng mô hình kinh tế Vi mô để phân tích cách ứng xử, hành vi của các chủ thể kinh tế khi họ theo đuổi mục đích của mình, như hành vi sản xuất và hành vi tiêu dùng, phân tích mối quan hệ giữa sản xuất và tiêu dùng, phân tích cân bằng thị trường.
        - *Phân tích hành vi sản xuất.*
        - *Phân tích tình huống tối ưu về mặt kinh tế của sản xuất.*
        - *Phân tích hành vi tiêu dùng.*
        - *Phân tích quan hệ giữa cung và cầu.*
        - *Phân tích cân bằng thị trường.*

5.1. Tổng quan

Sử dụng mô hình toán kinh tế trong nghiên cứu lựa chọn giải pháp tối ưu

- ***Sử dụng mô hình kinh tế Vĩ mô (Macro).***
  - Mô hình kinh tế vĩ mô phân tích mối quan hệ giữa các biến số kinh tế tổng quát (biến gộp) đặc trưng cho hoạt động Vĩ mô của nền kinh tế. Trong kinh tế thị trường người ta quan tâm đến ba khu vực: Thị trường hàng hoá - dịch vụ, thị trường tiền tệ và thị trường lao động.
    - *Phân tích tổng cung và tổng cầu.*
    - *Phân tích sự tác động của đầu tư đối với tổng sản phẩm của nền kinh tế quốc dân.*
    - *Phân tích vai trò của nhà nước trong quá trình điều tiết kinh tế thị trường.*
    - *Nghiên cứu sự tăng trưởng của các chỉ tiêu kinh tế quan trọng.*

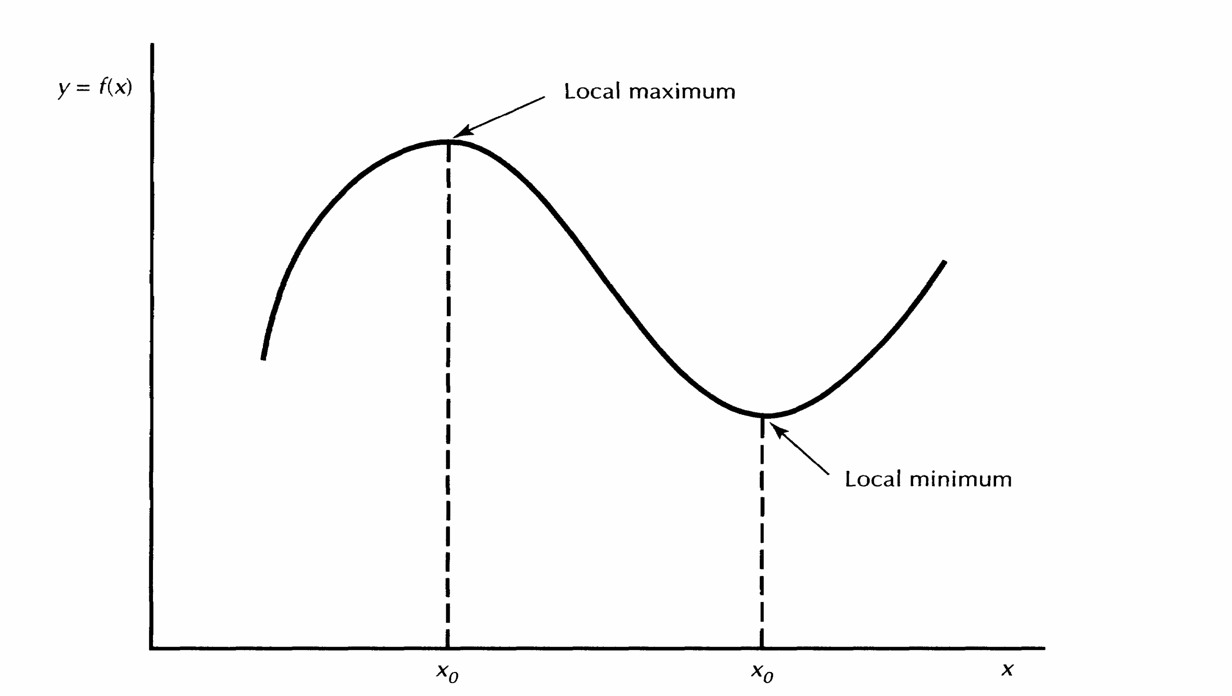
Cực trị của một hàm số

- *Phép thử đạo hàm bậc nhất* (*first derivative test* - hay *điều kiện bậc nhất*) dùng để xác định các điểm cực trị của một hàm (Xem hình bên). *dy/dx = y'= f'(x0) = 0* tại điểm cực đại hoặc cực tiểu và x_0 gọi là *cực trị* của hàm. Gồm 3 bước: (1) tìm đạo hàm, (2) đặt biểu thức bằng 0 và (3) tìm giá trị của x.
- Đạo hàm bậc 2 có được từ việc áp dụng quy tắc vi phân cho đạo hàm bậc nhất chứ không phải đối với hàm ban đầu. Kết quả vi phân đạo hàm bậc nhất cho ta đạo hàm bậc 2, có dạng $d^2y/dx^2 = y'' = f''(x)$.
- ***Ví dụ 1:***
  - Giả sử $y = f(x) = 3x^2$. Thì, $dy/dx = f'(x) = 6x$.
  - 6x là kết quả của đạo hàm bậc nhất. Đạo hàm bậc 2 là $d^2y/dx^2 = y'' = f''(x) = 6$.

5.2. Toán cao cấp và một số ứng dụng

Cực trị của một hàm số

- **ĐIỂM CỰC ĐẠI VÀ CỰC TIỂU**



- *Hàm $y = f(x)$ đạt cực đại tại điểm $x = x_0$ nào đó nếu $f'(x_0) = 0$ và $f''(x_0)$ âm.*
- *Tương tự như vậy, $f(x)$ đạt cực tiểu tại một điểm x_0 nào đó nếu $f'(x_0) = 0$ và $f''(x_0)$ dương.*
- *Nếu cả hàm bậc nhất và hàm bậc 2 đều bằng 0 thì ta chỉ có một điểm uốn chứ không có giá trị cực đại hoặc cực tiểu, tức là $f'(x_0) = 0$ và $f''(x_0) = 0$. Một ví dụ về hàm không có điểm cực đại và cực tiểu là $y = f(x) = x^3$.*

Cực trị của một hàm số

Ví dụ 2:

- Công ty cung cấp dịch vụ lau dọn Peruvian là nhà phân phối chủ yếu chất tẩy rửa quan trọng cho những người quét dọn khắp miền nam nước Mỹ. Chất tẩy rửa này dùng để tạo ra lớp bảo vệ bên ngoài cho các hầm làm lạnh suốt mùa hè có độ ẩm cao. Peruvian cung cấp chất tẩy trong những chiếc xe téc, và mỗi khách hàng phải mua ít nhất 100-gallon (1 gallon = 3.78 lít Mỹ). Giá tiền mỗi gallon là $12. Khách hàng mua khối lượng lớn hơn 100 gallon sẽ được chiết khấu $0.05 mỗi gallon. Phần trăm chiết khấu này chỉ áp dụng cho những lượng hàng lớn hơn mức tối thiểu; 100 gallon đầu tiên vẫn có giá là $12 mỗi gallon bất kể tổng số gallon được mua là bao nhiêu đi nữa.
- Hãy xác định lượng bán cho mỗi khách hàng để đạt doanh thu cực đại, tính doanh thu cực đại trên mỗi khách hàng.

5.2. Toán cao cấp và một số ứng dụng

Cực trị của một hàm số

Ví dụ 2:

- Giải:
  - Hàm toán học sau đây về tổng doanh thu từ mỗi khách hàng là:

    TR $= \$12(100) + [\$12 - \$0.05(x - 100)]\,(x - 100) =$

    $= -\,500 + 22x - 0.05x^2$
  - Đạo hàm của dạng hàm số này là: $d\mathrm{TR}/dx = 22 - 0.10x$
  - Ta cho đạo hàm bậc nhất bằng 0 và tìm được giá trị x, như sau:

    $22 - 0.10x = 0$; $\Rightarrow -\,0.10x = -\,22$; $\Rightarrow x = 220$.
  - *Do $d^2TR/dx^2 = -0.1$. Vì đạo hàm bậc 2 âm, giá trị cực trị của hàm ($x^* = 220$) là giá trị cực đại.*
  - Thay x bởi x* = 220 trong hàm gốc chúng ta có doanh thu cực đại như sau: $-\,500 + 22(220) - 0.05(220)^2 =$

    $= -\,500 + 4{,}840 - 0.05(48{,}400) = \$1{,}920$

Cực trị của một hàm số

Ví dụ 3:

- Mục tiêu của một hãng là tối đa hóa lợi nhuận. Để tìm ra sản lượng đầu ra có thể tối đa hóa lợi nhuận, chúng ta nên dùng phép tính vi phân. Giả sử ta có hàm tổng doanh thu (TR) và tổng chi phí (TC) sau đây:

  $TR(Q) = \$1{,}000Q - \$5Q^2$ và $TC(Q) = \$20{,}000 + \$200Q$.

- Khi đó Hàm lợi nhuận (π) là:

  $\pi\ (Q) = TR(Q) - TC(Q) = \$1{,}000Q - \$5Q^2 - (\$20{,}000 + \$200Q)$

  $= \$1{,}000Q - \$5Q^2 - \$20{,}000 - \$200Q = -\$20{,}000 + \$800Q - \$5Q^2$

- Lấy $d\pi/dQ = 0 \Rightarrow d\pi/dQ = \$800 - \$10Q = 0$; $Q^* = 80$ đơn vị

- Giá trị đạo hàm bậc 2 của hàm lợi nhuận là

  $d^2\pi/dQ^2 = -10 < 0$, cho biết $Q^* = 80$ là điểm tối đa hóa lợi nhuận.

5.2. Toán cao cấp và một số ứng dụng

Đạo hàm riêng và tối ưu hoá nhiều biến

- Nhiều hàm có chứa nhiều biến độc lập phức tạp. Khái niệm *tối ưu hóa nhiều biến* (*multivariate optimization)* và quá trình tối ưu hóa cho đẳng thức với nhiều biến quyết định là rất hữu ích.
- Để thực hiện, chúng ta phải tiến hành *vi phân riêng*.
- Các quy tắc vi phân riêng là giống nhau với điều kiện các biến độc lập không tham gia vào phép vi phân được xem như là những hằng số.

Đạo hàm riêng và tối ưu hoá nhiều biến

- **Ví dụ 4:** Để minh họa cho hàm tổng doanh thu là $TR = 2x^2y^2z$., trong đó x = chi phí quảng cáo trong giai đoạn trước, y = chi phí đi lại cho nhân viên kinh doanh, và z = là số hàng hóa mà đối thủ cạnh tranh bán ở thời điểm hiện tại. Giả thiết rằng ban quản lý cần biết giới hạn tối đa mà doanh thu thu được từ x có thể đạt tới (chi phí quảng cáo trong giai đoạn trước). Quy trình tìm giá trị cực đại như sau:
  - Vi phân riêng tương ứng với biến của thu nhập
  - Lấy đạo hàm riêng bằng 0 và tìm biến thu nhập
  - Xác định giá trị hàm gốc tại giá trị này để tìm cực trị

5.2. Toán cao cấp và một số ứng dụng

Đạo hàm riêng và tối ưu hoá nhiều biến

- Xem y và z là hằng số, hệ số đầy đủ của x^2 là $2y^2z$. Đạo hàm riêng tương ứng với x, biến số của thu nhập trong ví dụ này là: $\partial TR/\partial x = 4xy^2z$.
- Lấy đạo hàm riêng bằng 0 và tìm được x như sau:

$$4y^2zx = 0$$

$$x = 0.$$

- Do đó hàm doanh thu đạt cực trị khi x = 0.
- Đạo hàm riêng bậc 2 được xác định như sau: $\partial^2 TR/\partial x^2 = 4y^2z$. Vì đạo hàm riêng bậc 2 là dương nên giá trị cực trị của hàm là cực tiểu.

(Cần nhớ rằng quá trình vi phân coi y và z là hằng số, vì vậy không có kết luận gì về ảnh hưởng của những thay đổi trong các biến số đối với hàm doanh thu)

5.2. Toán cao cấp và một số ứng dụng

Tối ưu hoá bị ràng buộc

- Nhiều hãng phải đối mặt với những hạn chế trong các phương án quyết định. Chẳng hạn như hạn chế về nguồn lực (như tiền, thiết bị, năng lực sản xuất, nguyên liệu và nhân sự) sẵn có đối với hãng. Tối ưu hóa bị ràng buộc (Constrained optimization) là tối đa hóa lợi nhuận kèm theo những hạn chế trong sự sẵn có về nguồn lực, hoặc tối thiểu hóa chi phí kèm những yêu cầu tối thiểu cần được thỏa mãn. Những kỹ thuật như quy hoạch tuyến tính được dùng cho mục đích này.
- Vấn đề chung là tìm ra điểm cực trị của hàm f(x,y) tương ứng với các đẳng thức dạng: g(x,y) = 0
- Khi các ràng buộc dưới dạng *đẳng thức,* ta dùng các phương pháp tối ưu hóa cổ điển để tìm phương án tối ưu. Hai phương pháp thường dùng là: (1) Phương pháp thế và (2) Phương pháp nhân tử Lagrange.

Tối ưu hoá bị ràng buộc

- ***Phương pháp thế:***
  - Dùng phương pháp thế khi hàm mục tiêu chỉ phụ thuộc vào một biểu thức ràng buộc tương đối đơn giản. Bằng cách thế, chúng ta giảm được mức độ rắc rối của hàm mục tiêu. Phương pháp này gồm 2 bước: (1) tìm ra được một trong nhiều biến quyết định thỏa mãn nhất sau đó (2) thay giá trị của biến này vào hàm mục tiêu. Quá trình này chuyển từ hàm ban đầu sang hàm tối ưu hóa *không bị ràng buộc* để áp dụng được phép tính vi phân được áp dụng nhằm tìm ra phương án tối ưu.
  - Hạn chế của phương pháp thế đó là nó chỉ thực hiện được khi chỉ có một ràng buộc và chỉ có thể giải ra một biến. Từ 2 điều kiện trở lên và/hoặc có cấu trúc ràng buộc phức tạp thì sử dụng phương pháp nhân tử Lagrange.

Tối ưu hoá bị ràng buộc

- ***Phương pháp thế:***
  - **<u>Ví dụ 5:</u>** Giả sử một hãng sản xuất với 2 dây chuyền lắp ráp tự động và hoạt động với hàm tổng chi phí có dạng TC(x, y) = $3x^2 + 6y^2 - xy$, trong đó x = sản lượng đầu ra của dây chuyền thứ nhất và y = sản lượng đầu ra của dây chuyền thứ hai. Các nhà quản lý cần phải quyết định phương pháp kết hợp x và y sao cho tốn ít chi phí nhất, với điều kiện rằng tổng đầu ra phải là 20 đơn vị.
  - Vấn đề tối ưu hóa với điều kiện ràng buộc ở trên có thể được giải quyết như sau:
    - *Tối thiểu hóa* $TC(x, y) = 3x^2 + 6y^2 - xy$
    - *Ràng buộc:* $x + y = 20$

5.2. Toán cao cấp và một số ứng dụng

Tối ưu hoá bị ràng buộc

- ***Phương pháp thế:***

  Giải:

  - Giải biểu thức ràng buộc để tìm x, có x = 20 – y và thế vào hàm mục tiêu. $TC(x, y) = T(y) = 3(20 - y)^2 + 6y^2 - (20 - y)y = 3(400 - 40y + y^2) + 6y^2 - 20y + y^2 = 1{,}200 - 120y + 3y^2 + 6y^2 - 20y + y^2 = 1{,}200 - 140y + 10y^2$
  - Lấy đạo hàm của hàm thế đã giản lược và cho nó bằng 0, ta có:

    $dTC/dy = -140 + 20y = 0$; y = 7 đơn vị.

    Thế ngược trở lại vào x: x = 20 - y = 20 - 7 = 13 đơn vị.
  - Do vậy x = 13 và y = 7 là phương án tối ưu cho vấn đề tối thiểu hóa chi phí bị ràng buộc.

5.2. Toán cao cấp và một số ứng dụng

Tối ưu hoá bị ràng buộc

- ***Phương pháp nhân tử Lagrange:***
  - Một phương pháp để giải quyết các vấn đề *tối ưu hóa bị ràng buộc* mà trong đó có ràng buộc đối với hàm mục tiêu ban đầu (làm cho hàm mục tiêu bằng 0 khi thỏa mãn). Hàm mục tiêu mới đã *thêm* ràng buộc được gọi là *hàm Lagrange,* sẽ tạo ra một bài toán tối ưu hóa không bị ràng buộc có cấu trúc như sau:

    $L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$
  - Hệ số của đẳng thức ràng buộc g(x,y), λ (đọc là lamda), gọi là *nhân tử Lagrange.* Vì đẳng thức ràng buộc bằng 0 nên khi thêm $\lambda g(x, y)$ vào hàm mục tiêu f(x, y) không làm thay đổi giá trị của hàm.
  - Biến giả này cho biết sự thay đổi cận biên trong giá trị của hàm mục tiêu có được từ sự thay đổi của một đơn vị trong giá trị của ràng buộc. Sử dụng phương pháp Lagrange khi (1) không dùng được phương pháp thế và (2) khi có nhiều ràng buộc.

5.2. Toán cao cấp và một số ứng dụng

Tối ưu hoá bị ràng buộc

- ***Phương pháp nhân tử Lagrange:***
  - **<u>Từ ví dụ 5,</u>** ta có ràng buộc x + y = 20. Trước hết, cần biến đổi sao cho đẳng thức có một vế bằng 0, g(x, y) = 0.

    Khi đó x + y = 20 có dạng 20 - x – y = 0.
  - Tiếp theo, chúng ta xác định biến giả λ và xây dựng Hàm Lagrange (L):

    $L(x, y, \lambda) = TC(x, y) + \lambda g(x, y) = 3x^2 + 6y^2 - xy + \lambda(20 - x - y)$
  - Vì $L(x, y, \lambda)$ là một hàm với 3 biến quyết định nên để tối thiểu hóa hàm này cần:
    - *Vi phân riêng theo mỗi biến*
    - *Cho các biểu thức đạo hàm riêng bằng 0*
    - *Giải các đẳng thức để tìm các giá trị x, y và λ .*

5.3. Mô hình qui hoạch tuyến tính

Giới thiệu qui hoạch tuyến tính

- Quy hoạch tuyến tính (Linear programming - LP) là một thuật toán nhằm tìm ra phương án tối ưu (hoặc kế hoạch tối ưu) từ vô số các phương án quyết định. Phương án tối ưu là phương án thỏa mãn được các mục tiêu đề ra của một hãng, phụ thuộc vào các hạn chế và các ràng buộc. Quyết định tối ưu mang lại hiệu quả cao nhất, lãi gộp (Contribution Margin-CM) cao nhất, hay doanh thu, hay chi phí thấp nhất. Mô hình LP gồm 2 thành phần:
  - ***Hàm mục tiêu*:** Xác định mục tiêu cụ thể phải đạt tới.
  - ***Các ràng buộc:*** Các ràng buộc dưới dạng các hạn chế về sự sẵn có của nguồn lực hay thoả mãn các yêu cầu tối thiểu. Như tên gọi quy hoạch tuyến tính, cả hàm mục tiêu và các ràng buộc phải dưới dạng *tuyến tính*.

5.3. Mô hình qui hoạch tuyến tính

Giới thiệu qui hoạch tuyến tính

- LP có nhiều ứng dụng. Bao gồm:
  - Lựa chọn kết hợp đầu vào có chi phí thấp nhất cho sản phẩm sản xuất ra.
  - Xác định ngân sách tối ưu.
  - Quyết định danh mục đầu tư tối ưu (hay phân bổ tài sản).
  - Phân bổ ngân sách quảng cáo cho các phương tiện thông tin.
  - Quyết định phương thức vận chuyển có chi phí thấp nhất.
  - Kết hợp khí đốt.
  - Phân bố nhân lực tối ưu.
  - Lựa chọn vị trí đặt nhà xưởng phù hợp nhất.

5.3. Mô hình qui hoạch tuyến tính

Giới thiệu qui hoạch tuyến tính

<u>Ví dụ 6:</u>

- Công ty sản xuất đồ nội thất Omni sản xuất 2 sản phẩm: bàn và ghế. Cả 2 sản phẩm cần thời gian để được xử lý trong 2 bộ phận: Bộ phận mộc và bộ phận sơn. Dữ liệu về hai sản phẩm này như sau.

Xử lý	Bàn (Chiếc)	Ghế (Chiếc)	Số giờ sẵn có
Hiệu quả	\$7	\$5	
Phần mộc	3 hrs	4 hrs	2400
Phần sơn	2 hrs	1 hr	1000

- Ràng buộc bổ sung: Sản xuất không quá 450 ghế và ít nhất 100 bàn
- Công ty muốn tìm được cách kết hợp 2 loại sản phẩm này sao cho có lợi nhất.

Giới thiệu qui hoạch tuyến tính

Ví dụ 6:

- *Bước 1:* Xác định các biến quyết định như sau:

  x_1 = Số lượng bàn

  x_2 = Số lượng ghế

- *Bước 2:* Hàm mục tiêu để tối đa hoá hiệu quả (Z) được biểu diễn dưới đây:

  $Z = 7x_1 + 5x_2$

  Sau đó lập công thức các ràng buộc như là các bất đẳng thức:

  $3x_1 + 4x_2 \leq 2400$ (Ràng buộc công đoạn mộc)

  $2x_1 + 1x_2 \leq 1000$ (Ràng buộc công đoạn sơn)

  $x_1 \geq 100$; $x_2 \leq 450$

  Thêm vào đó, ẩn trong bất kỳ công thức LP nào phải có điều kiện để làm cho x_1 và x_2 không âm, tức là: $x_1, x_2 \geq 0$.

5.3. Mô hình qui hoạch tuyến tính

Các phương pháp tính toán

- Có nhiều phương pháp để tính toán LP bao gồm:
  - Phương pháp đơn hình
  - Phương pháp đồ thị
- Phương pháp đơn hình là phương pháp được sử dụng giải bài toán LP. Nó là một thuật toán, một phương pháp tính toán lặp đi lặp lại, từ phương án này tới phương án khác cho đến khi đạt được lời giải tốt nhất. Mặc dù vậy, hiện nay với sự trợ giúp của rất nhiều phần mềm ứng dụng (kể cả Excel), chúng ta dễ dàng tìm được phương án tối ưu mà không cần phải giải bằng tay như trước.
- Phương pháp đồ thị dễ sử dụng hơn nhưng chỉ đối với các trường hợp LP có 2 biến quyết định.

5.3. Mô hình qui hoạch tuyến tính

Các phương pháp tính toán- Phương pháp đồ thị

Phương pháp đồ thị gồm các bước sau đây:

- *Bước 1*: Đưa bất đẳng thức về dạng đẳng thức.
- *Bước 2*: Minh họa bằng đồ thị các đẳng thức. Để minh họa:
  - Đặt một biến bằng 0 và tìm giá trị biến còn lại và nối 2 giá trị trên đồ thị,
  - Đánh dấu các điểm trên 2 trục và kết nối với nhau thành 1 đường thẳng.
- *Bước 3:* Xác định phần thỏa mãn của các đẳng thức bằng cách đánh bóng. Lặp lại các bước từ 1-3 đối với mỗi ràng buộc.
- *Bước 4:* Sau hết, xác định tập phương án tức là đánh dấu các vùng chứa các phương án thoả mãn tất cả các ràng buộc.
- *Bước 5:* Giải đồng thời các ràng buộc (thể hiện dưới dạng các đẳng thức) để tìm ra điểm cận biên.
- *Bước 6:* Xác định hiệu quả hoặc lãi gộp tại tất cả các đỉnh trong miền khả thi.

5.3. Mô hình qui hoạch tuyến tính

Các phương pháp tính toán- Phương pháp đồ thị

CHÚ Ý:

Tập phương án là những giá trị của biến quyết định thoả mãn đồng thời các ràng buộc. Chúng được tìm thấy phía trên và bên trong miền khả thi. Phương pháp đồ thị dựa vào 2 đặc điểm quan trọng của LP:

1. Phương án tối ưu nằm ở đường biên của vùng khả thi, có nghĩa là có thể bỏ qua các điểm bên trong vùng khả thi (rất nhiều điểm) khi tìm kiếm phương án tối ưu.
2. Phương án tối ưu nằm ở 1 trong các đỉnh của miền tối ưu (các phương án khả thi cơ bản)

5.3. Mô hình qui hoạch tuyến tính

Các phương pháp tính toán- Phương pháp đồ thị

Giải:

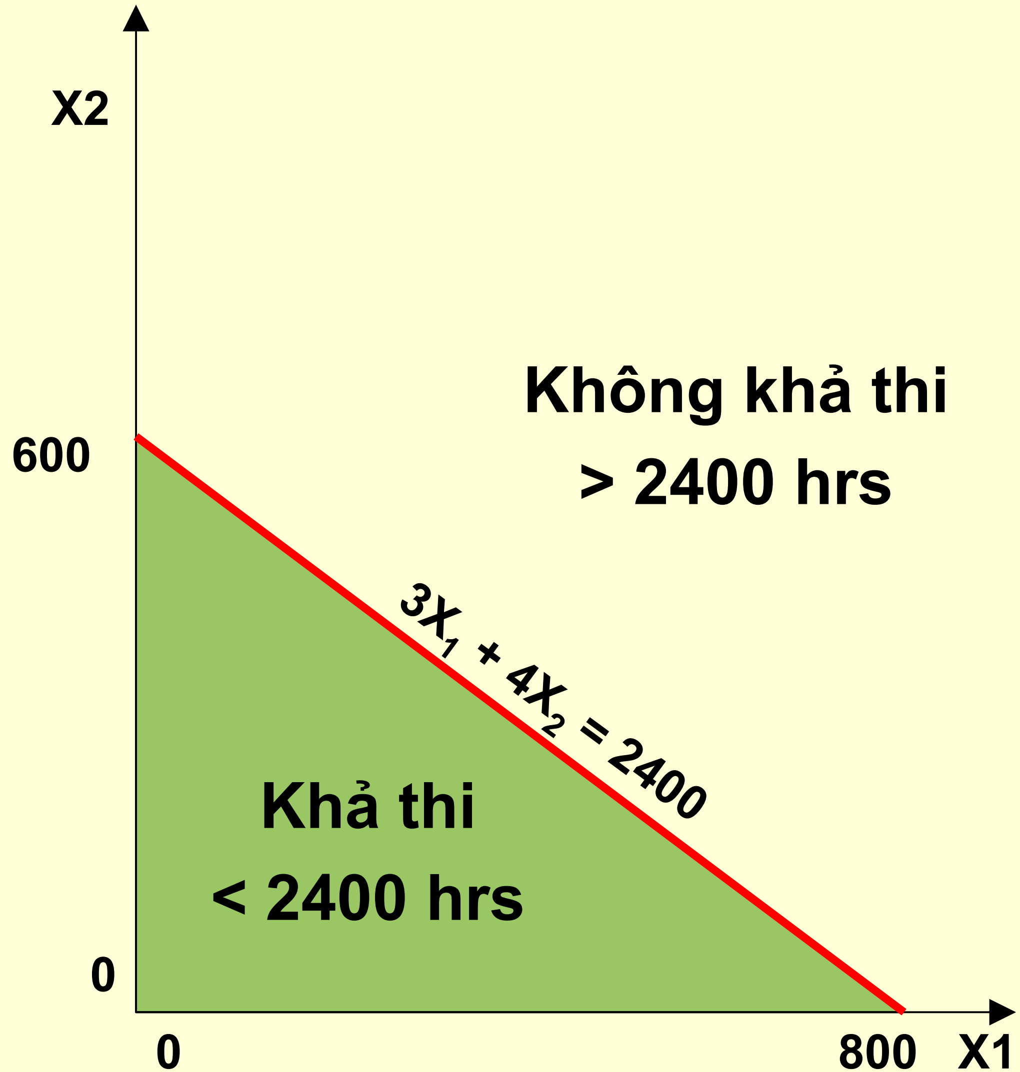
Đường giới hạn <u>công đoạn mộc:</u>

$3X_1 + 4X_2 = 2400$

<u>Chặn:</u>

$(X_1 = 0, X_2 = 600)$

$(X_1 = 800, X_2 = 0)$



5.3. Mô hình qui hoạch tuyến tính

Các phương pháp tính toán- Phương pháp đồ thị

Giải:

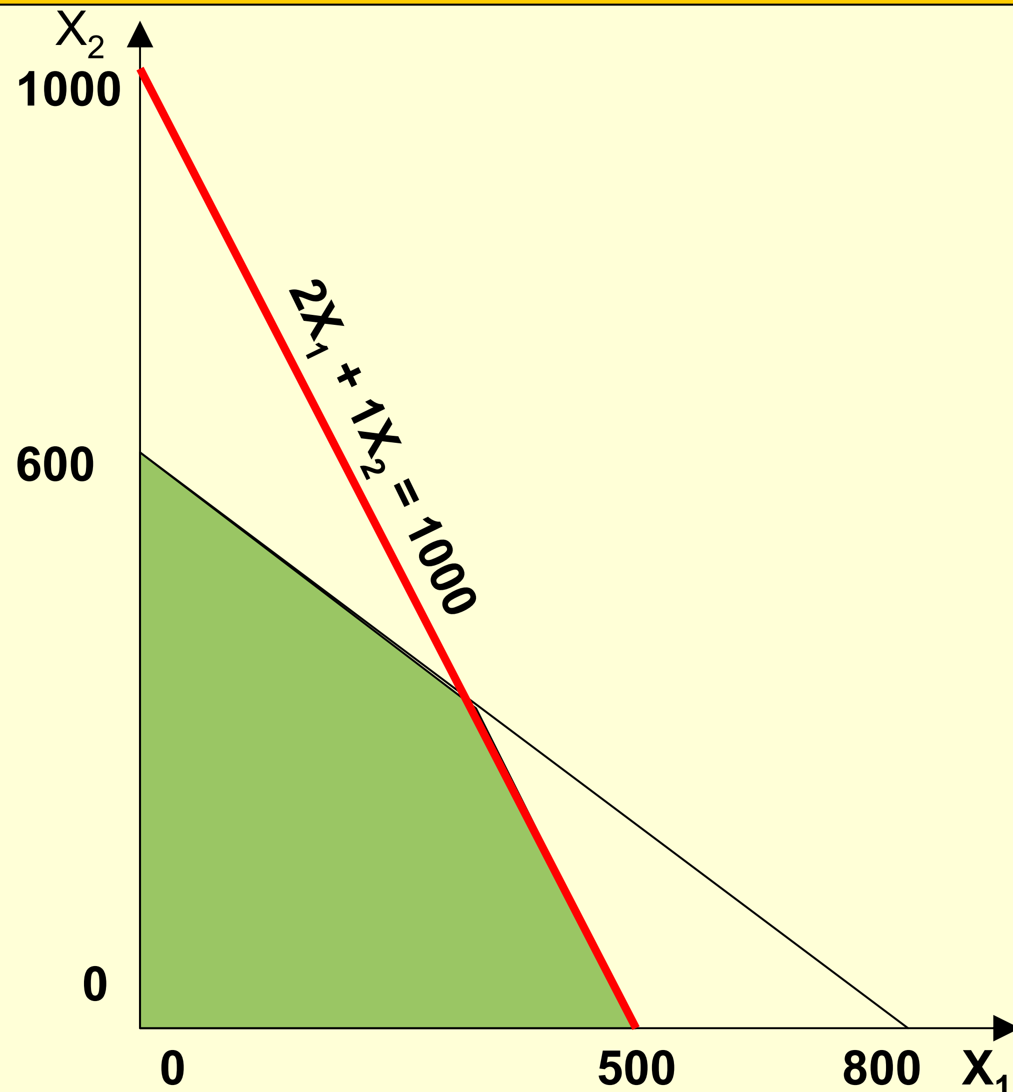
Đường giới hạn
<u>công đoạn sơn:</u>

$2X_1 + 1X_2 = 1000$

<u>Chặn:</u>

$(X_1 = 0, X_2 = 1000)$

$(X_1 = 500, X_2 = 0)$



5.3. Mô hình qui hoạch tuyến tính

Các phương pháp tính toán- Phương pháp đồ thị

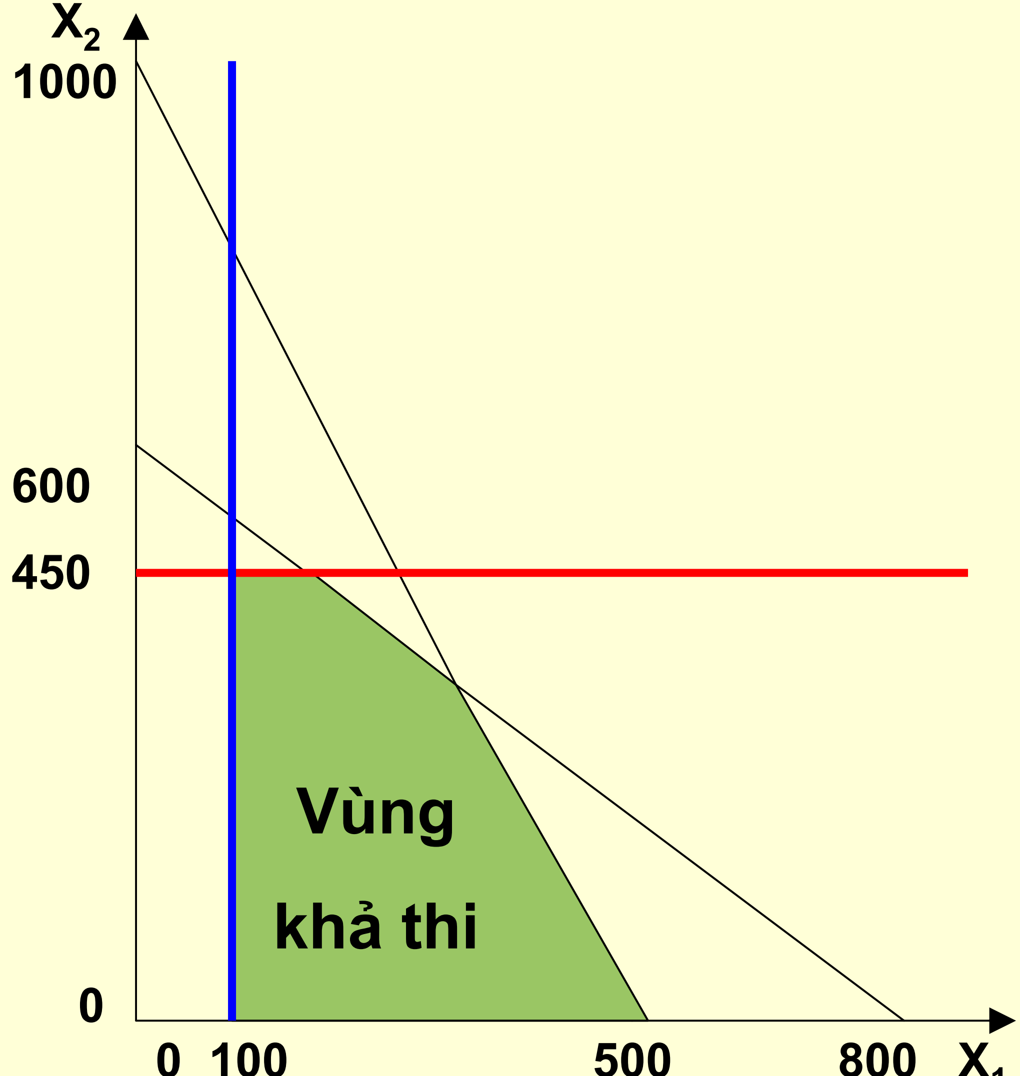
Giải:

Đường tối đa "ghế"

$X_2 = 450$

Đường tối thiểu "bàn"

$X_1 = 100$



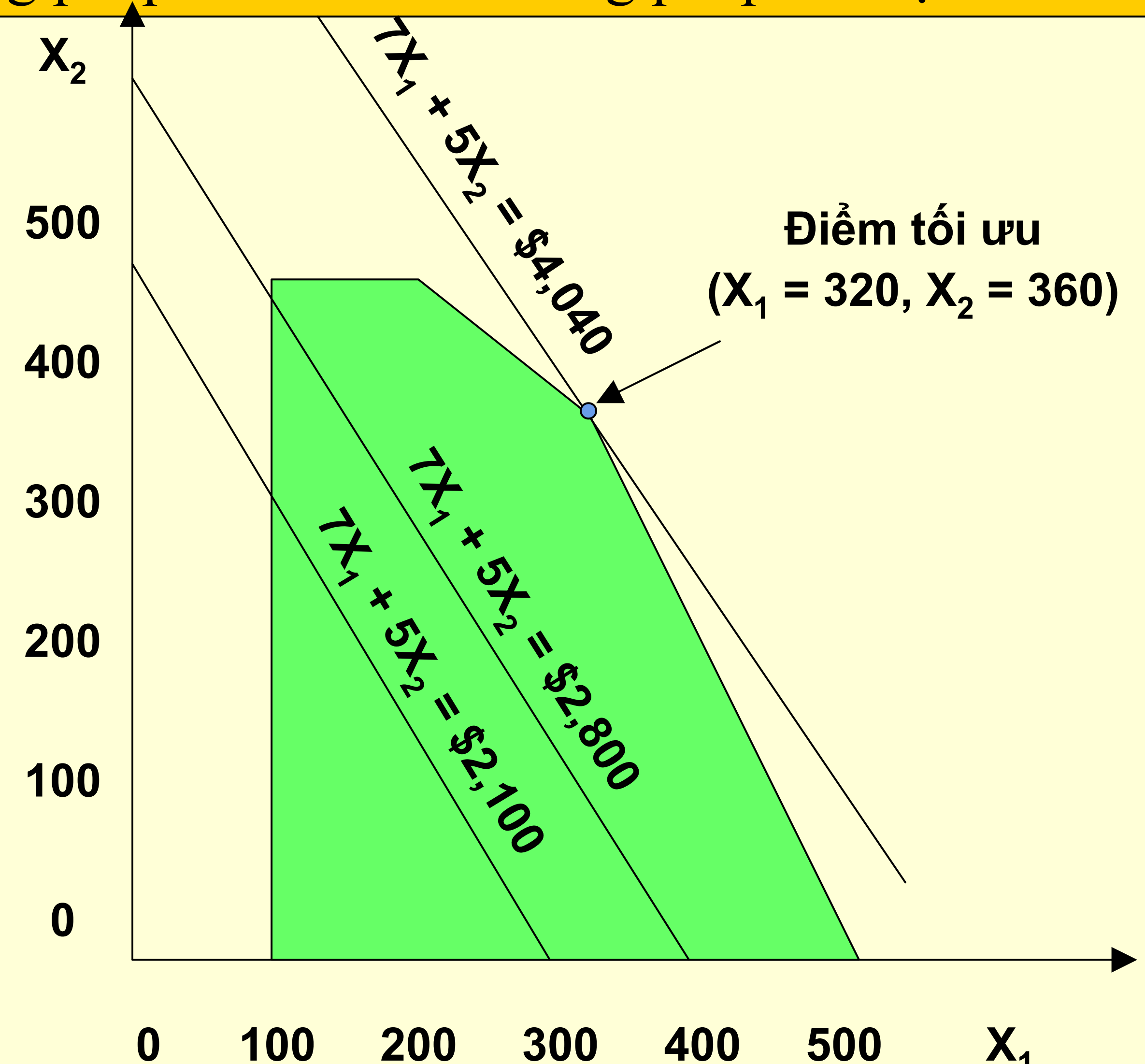
5.3. Mô hình qui hoạch tuyến tính

Các phương pháp tính toán- Phương pháp đồ thị

Giải:

<u>Đường hàm mục tiêu</u>

$7X_1 + 5X_2$ = Lợi nhuận



5.3. Mô hình qui hoạch tuyến tính

Phân tích độ nhạy cho bài toán quy hoạch và ứng dụng Solver

- Phân tích độ nhạy là nghiên cứu sự thay đổi của những hệ số trong bài toán qui hoạch tuyến tính ảnh hưởng đến phương án tối ưu.
- Dùng phân tích độ nhạy, chúng ta có thể trả lời những câu hỏi sau:
  - Hệ số trong hàm mục tiêu thay đổi sẽ ảnh hưởng như thế nào đến phương án tối ưu?
  - Giá trị của vế phải của các ràng buộc thay đổi sẽ ảnh hưởng như thế nào đến phương án tối ưu?
  - Trong nguồn lực sản xuất, nhân tố nào quan trọng hơn?.
- Phân tích độ nhạy thường được gọi là phân tích hậu tối ưu. Phân tích độ nhạy rất quan trọng trong việc ra quyết định vì các bài toán tồn tại trong môi trường thay đổi. Phân tích độ nhạy cung cấp những thông tin cần thiết ứng với những thay đổi đó.

5.3. Mô hình qui hoạch tuyến tính

Phân tích độ nhạy cho bài toán quy hoạch và ứng dụng Solver

- Chúng ta có thể thực hiện phân tích độ nhạy bằng phương pháp đồ thị hay bằng bảng đơn hình. Theo hướng ứng dụng, chúng ta không đi sâu vào phân tích bằng bảng đơn hình, chúng ta sẽ thực hiện qua Excel solver.
- Nhằm triển khai ý tưởng thực hiện phân tích độ nhạy, chúng ta xem xét bài toán tối ưu như sau: Công ty Galaxy sản xuất 2 loại sản phẩm là SD và ZD. Nguyên liệu sử dụng là 1 loại nhựa đặc biệt.
    - Định mức chi phí nguyên liệu và nhân công cho việc sản xuất 2 sản phẩm như sau:
        - *SD cần 2 cân nhựa và 3 phút giờ công lao động.*
        - *ZD cần 1 cân nhựa và 4 phút giờ công lao động.*
    - Trong đó giới hạn về nguồn lực là: 1000 cân nhựa và 40 giờ làm việc mỗi tuần.

Phân tích độ nhạy cho bài toán quy hoạch và ứng dụng Solver

- Yêu cầu từ bộ phận Marketing:
  - Tổng số lượng sản xuất không quá 700 tá.
  - Số lượng SD không vượt quá số lượng ZD là 350 tá.
- Dự kiến: Lợi nhuận thu được là $8/ tá SD, $5/ tá ZD.
- Kế hoạch sản xuất hiện tại là:
  - SD = 450 tá
  - ZD = 100 tá
  - Lợi nhuận = $4100/ tuần

8(450) + 5(100)

Ban giám đốc đang tìm kiếm phương án sản xuất nhằm gia tăng lợi nhuận cho Công ty

5.3. Mô hình qui hoạch tuyến tính

Phân tích độ nhạy cho bài toán quy hoạch và ứng dụng Solver

Ứng dụng mô hình qui hoạch tuyến tính

- Biến quyết định:
  - X_1 = Số lượng sản xuất sản phẩm SD (tá/tuần)
  - X_2 = Số lượng sản xuất sản phẩm ZD (tá/tuần) .
- Hàm mục tiêu: Tối đa hoá lợi nhuận/ tuần

Phân tích độ nhạy cho bài toán quy hoạch và ứng dụng Solver

Ứng dụng mô hình qui hoạch tuyến tính

Max $8X_1 + 5X_2$ (Lợi nhuận tuần)

Các ràng buộc

$2X_1 + 1X_2 \leq 1000$ (Nhựa)

$3X_1 + 4X_2 \leq 2400$ (Thời gian sản xuất)

$X_1 + X_2 \leq 700$ (Tổng số lượng sản xuất)

$X_1 - X_2 \leq 350$ (Mix)

$X_j >= 0,\ j = 1,2$ (Không âm)

Phân tích độ nhạy cho bài toán quy hoạch và ứng dụng Solver

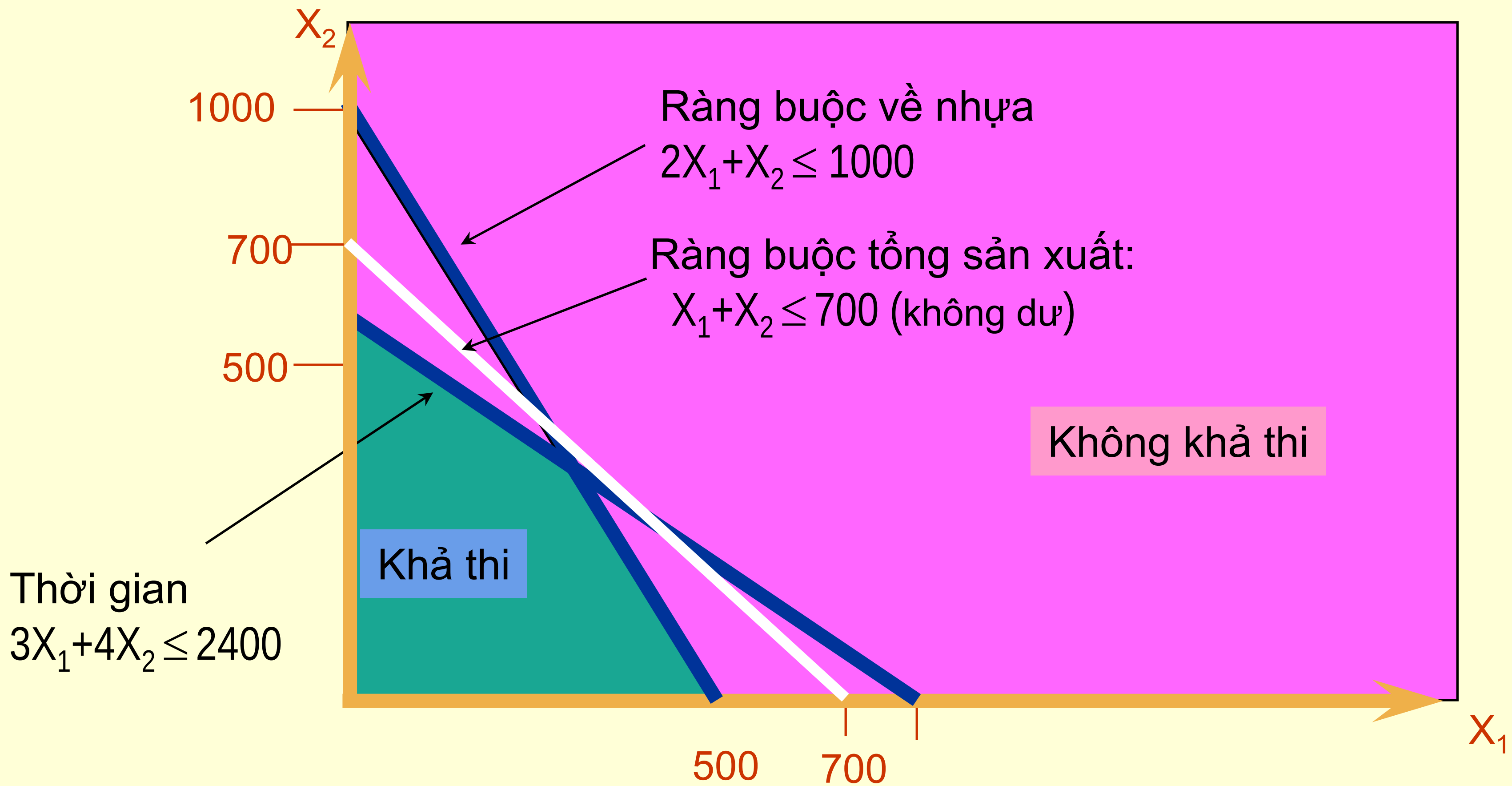
Ứng dụng mô hình qui hoạch tuyến tính- Giải bằng đồ thị

- Sử dụng đồ thị để mô tả các ràng buộc, hàm mục tiêu và miền khả thi.

Phân tích độ nhạy cho bài toán quy hoạch và ứng dụng Solver

Ứng dụng mô hình qui hoạch tuyến tính- Giải bằng đồ thị

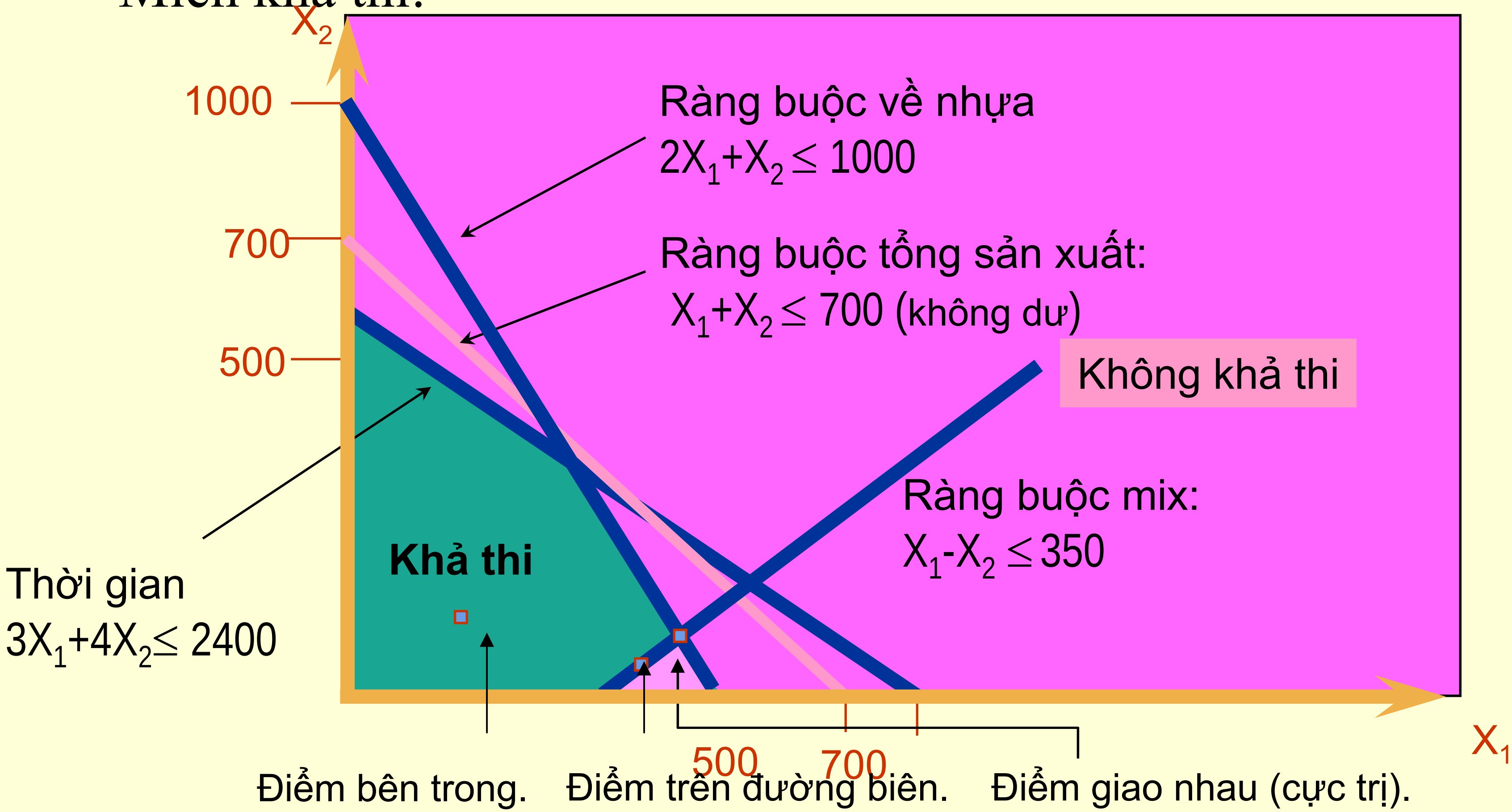
- Miền khả thi:



Phân tích độ nhạy cho bài toán quy hoạch và ứng dụng Solver

Ứng dụng mô hình qui hoạch tuyến tính- Giải bằng đồ thị

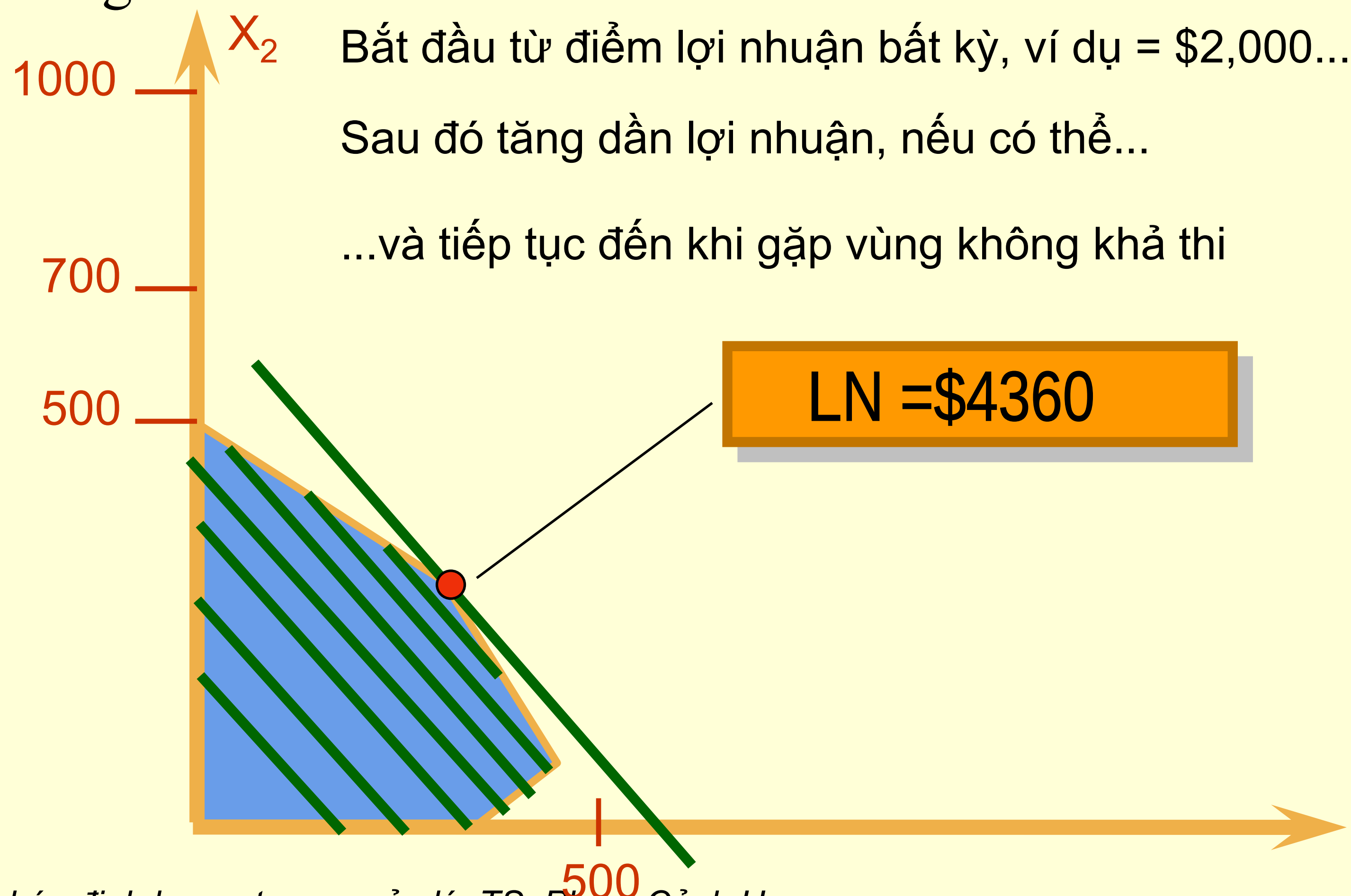
- Miền khả thi:



Phân tích độ nhạy cho bài toán quy hoạch và ứng dụng Solver

Ứng dụng mô hình qui hoạch tuyến tính- Giải bằng đồ thị

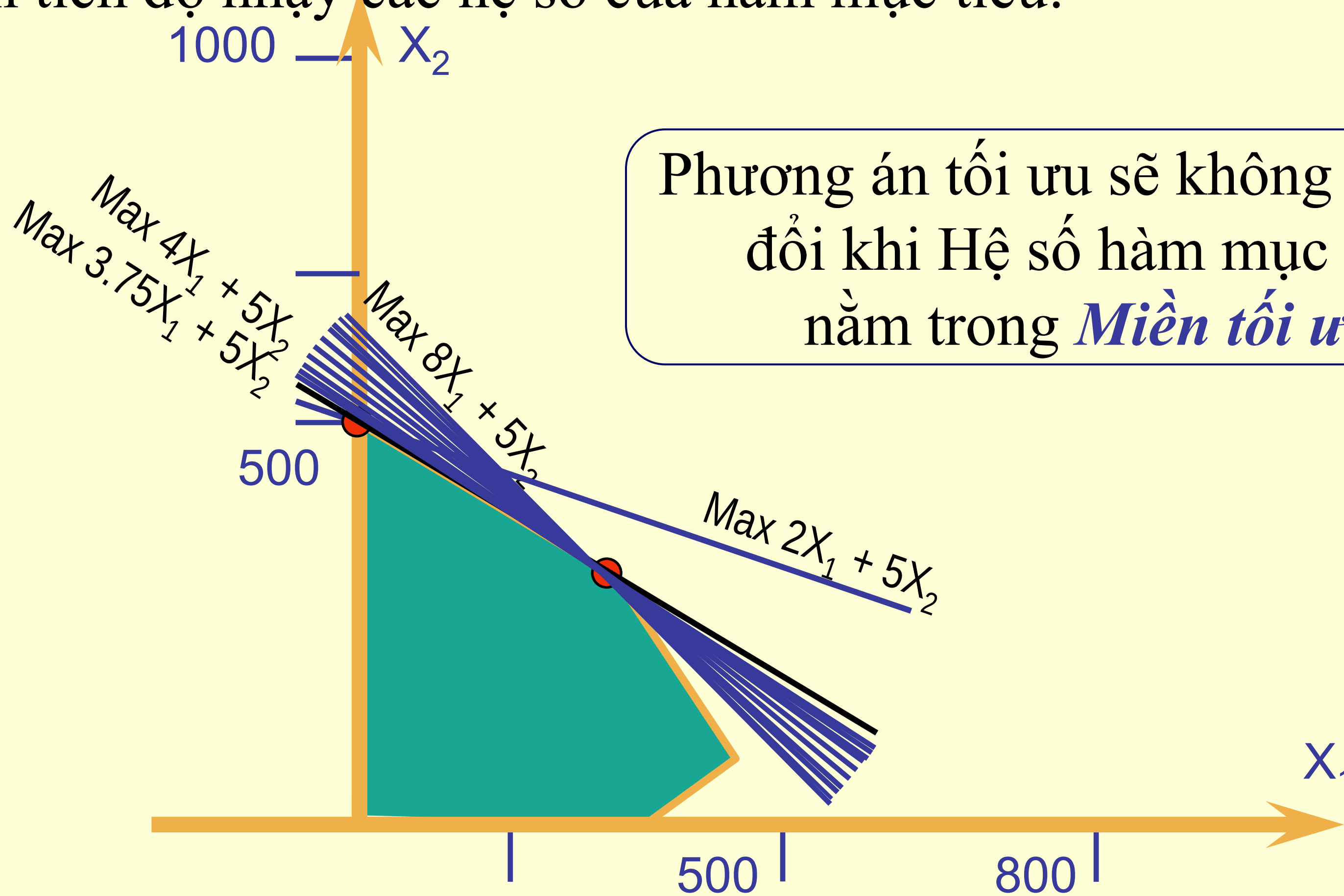
- Phương án tối ưu:



Phân tích độ nhạy cho bài toán quy hoạch và ứng dụng Solver

Ứng dụng mô hình qui hoạch tuyến tính- Phân tích độ nhạy

- Phân tích độ nhạy các hệ số của hàm mục tiêu:

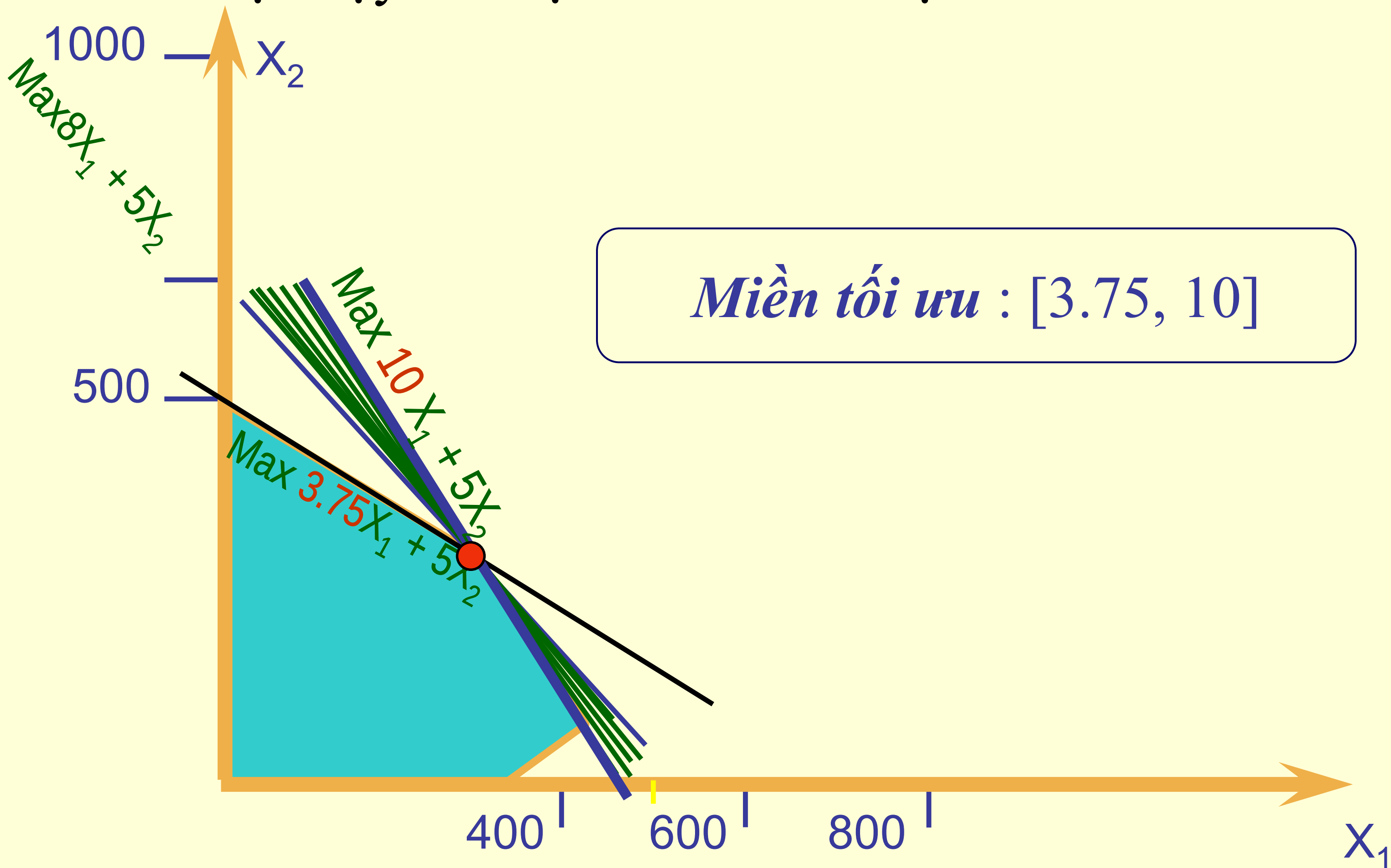


Phương án tối ưu sẽ không thay đổi khi Hệ số hàm mục tiêu nằm trong ***Miền tối ưu***

Phân tích độ nhạy cho bài toán quy hoạch và ứng dụng Solver

Ứng dụng mô hình qui hoạch tuyến tính- Phân tích độ nhạy

- Phân tích độ nhạy các hệ số của hàm mục tiêu:



Phân tích độ nhạy cho bài toán quy hoạch và ứng dụng Solver

Ứng dụng mô hình qui hoạch tuyến tính- Phân tích độ nhạy

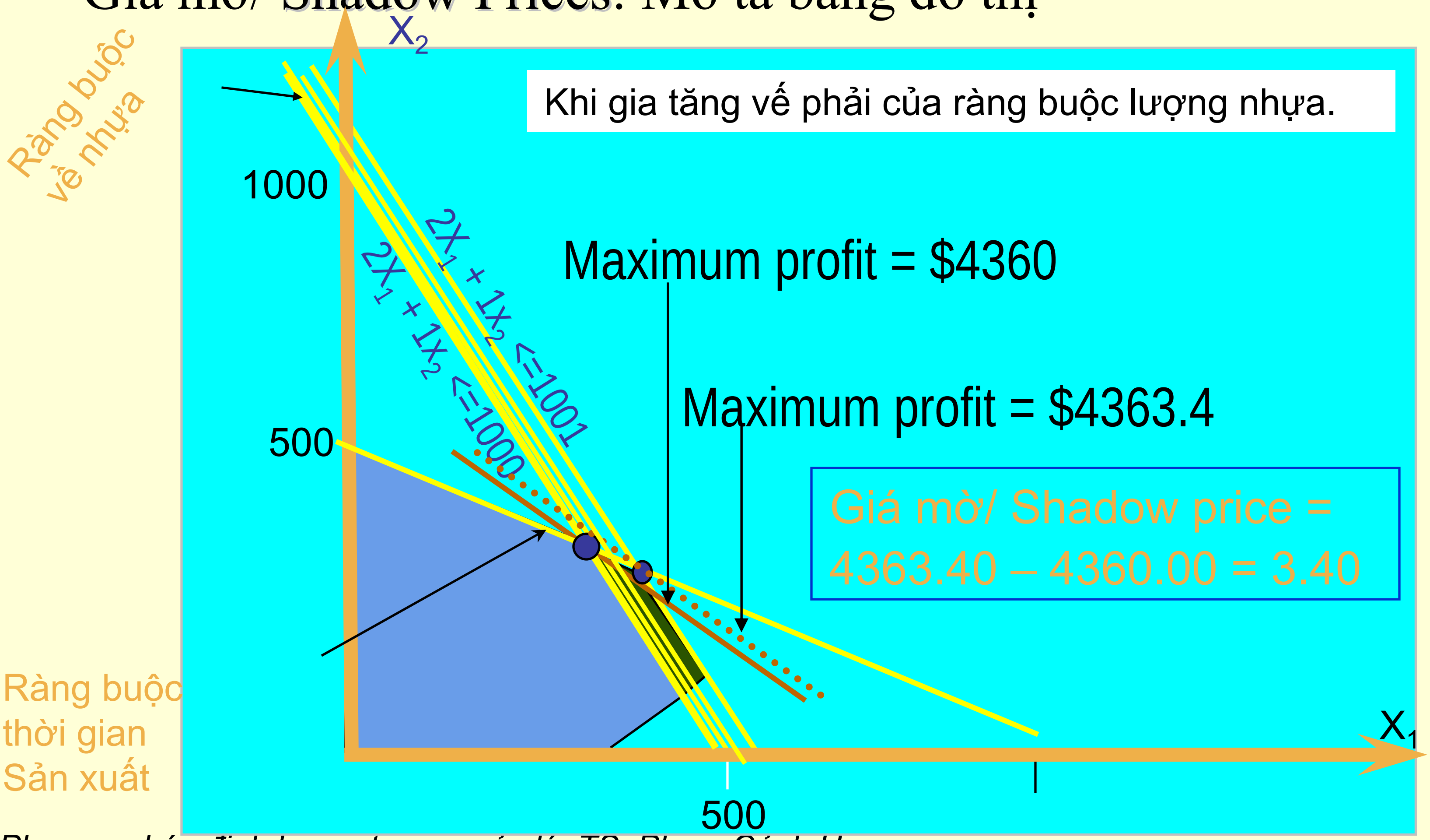
- Giá mờ/ Shadow Prices:

> Giả sử không có những thay đổi nào của các thông số đầu vào, giá trị thay đổi của hàm mục tiêu khi gia tăng một đơn vị (phía phải) của ràng buộc được gọi là "giá mờ"

Phân tích độ nhạy cho bài toán quy hoạch và ứng dụng Solver

Ứng dụng mô hình qui hoạch tuyến tính- Phân tích độ nhạy

- Giá mờ/ Shadow Prices: Mô tả bằng đồ thị



Phân tích độ nhạy cho bài toán quy hoạch và ứng dụng Solver

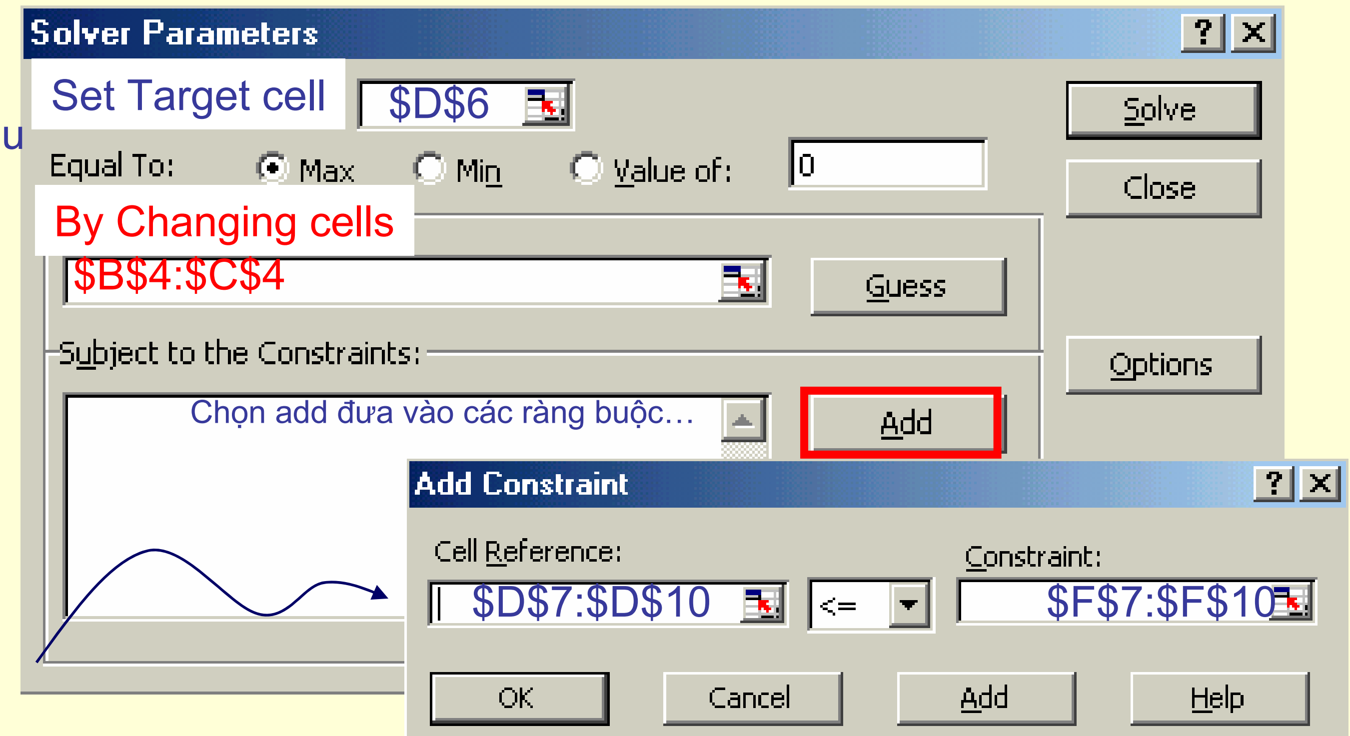
Sử dụng Excel Solver tìm phương án tối ưu

- Excel: [Galaxy.xls]
- Chọn Solver, ta thấy xuất hiện hộp thoại

Đây là ô chứa giá trị hàm mục tiêu

Vùng chứa biến Quyết định

Nhập vào các Ràng buộc.



5.3. Mô hình qui hoạch tuyến tính

Phân tích độ nhạy cho bài toán quy hoạch và ứng dụng Solver

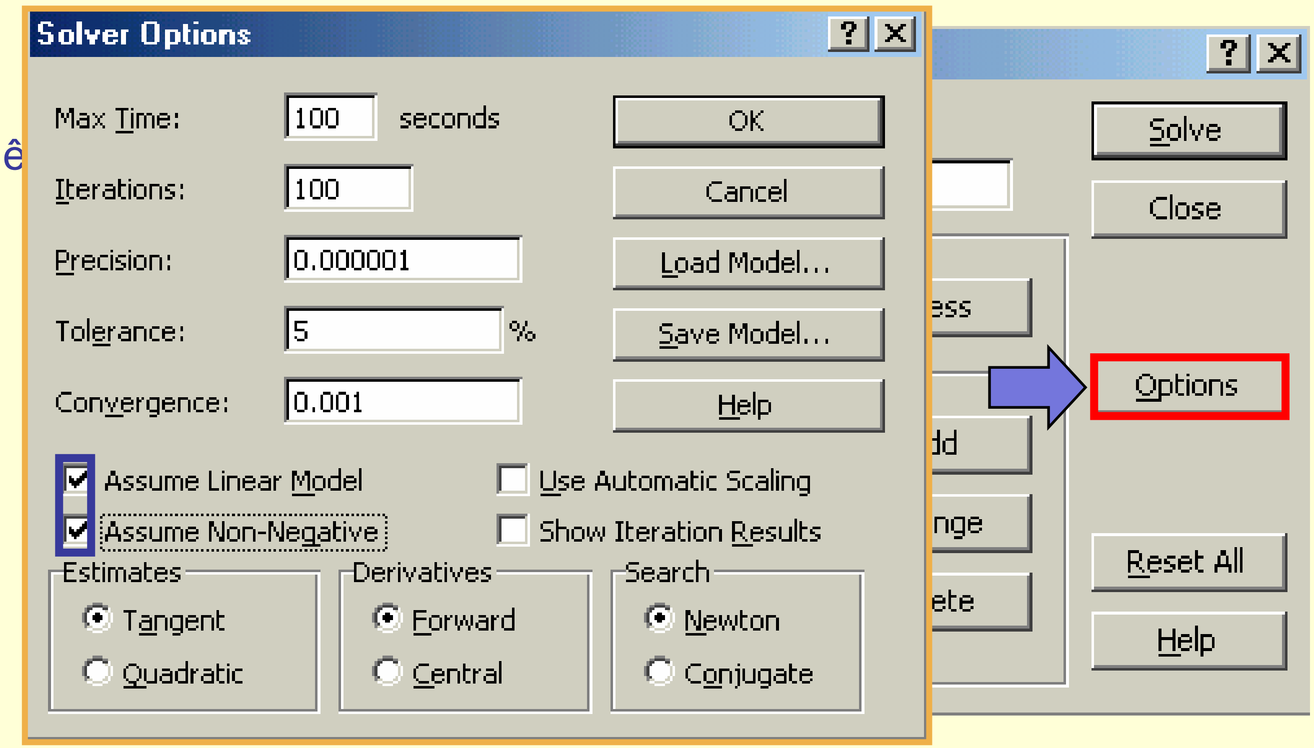
Sử dụng Excel Solver tìm phương án tối ưu

- Excel: [Galaxy.xls]
- Chọn Solver, ta thấy xuất hiện hộp thoại

Đây là ô chứa
giá trị hàm mục tiê

Vùng chứa biến
Quyết định

Chọn 'Options'
Và chọn 'Linear
Programming' &
'Non-negative'.

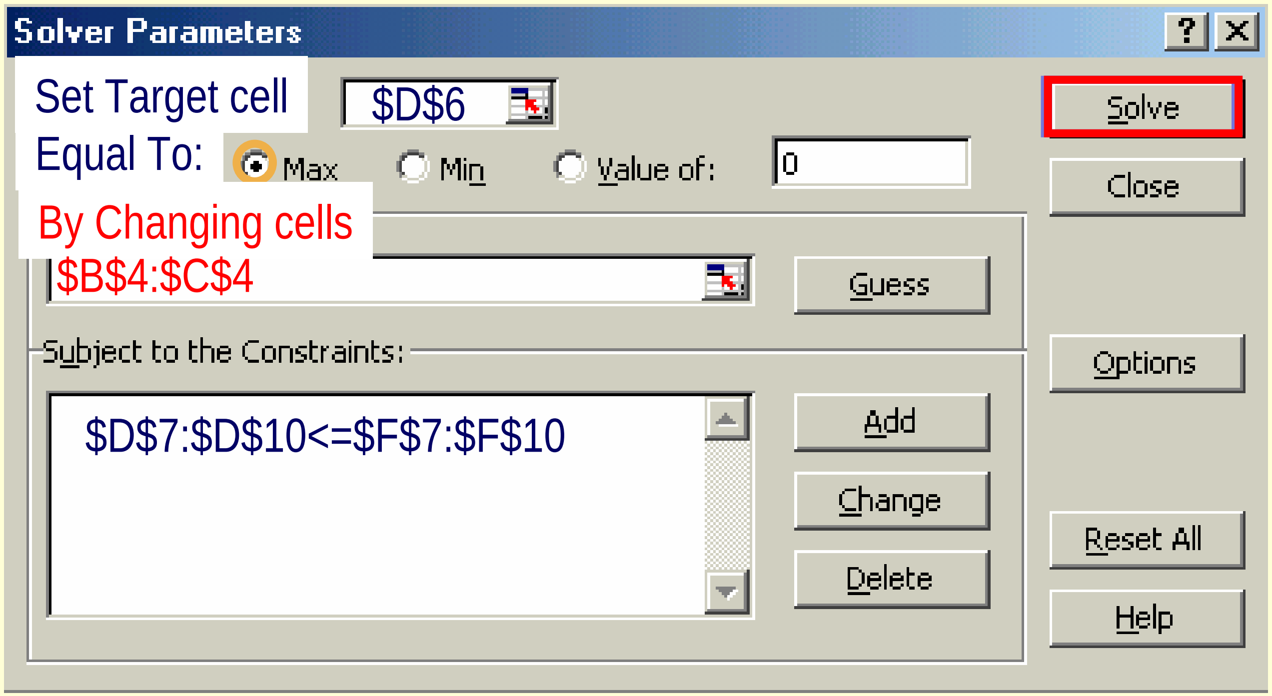


199

Phân tích độ nhạy cho bài toán quy hoạch và ứng dụng Solver

Sử dụng Excel Solver tìm phương án tối ưu

- Excel: Galaxy.xls
- Chọn Solver, ta thấy xuất hiện hộp thoại



5.3. Mô hình qui hoạch tuyến tính

Phân tích độ nhạy cho bài toán quy hoạch và ứng dụng Solver

Sử dụng Excel Solver tìm phương án tối ưu

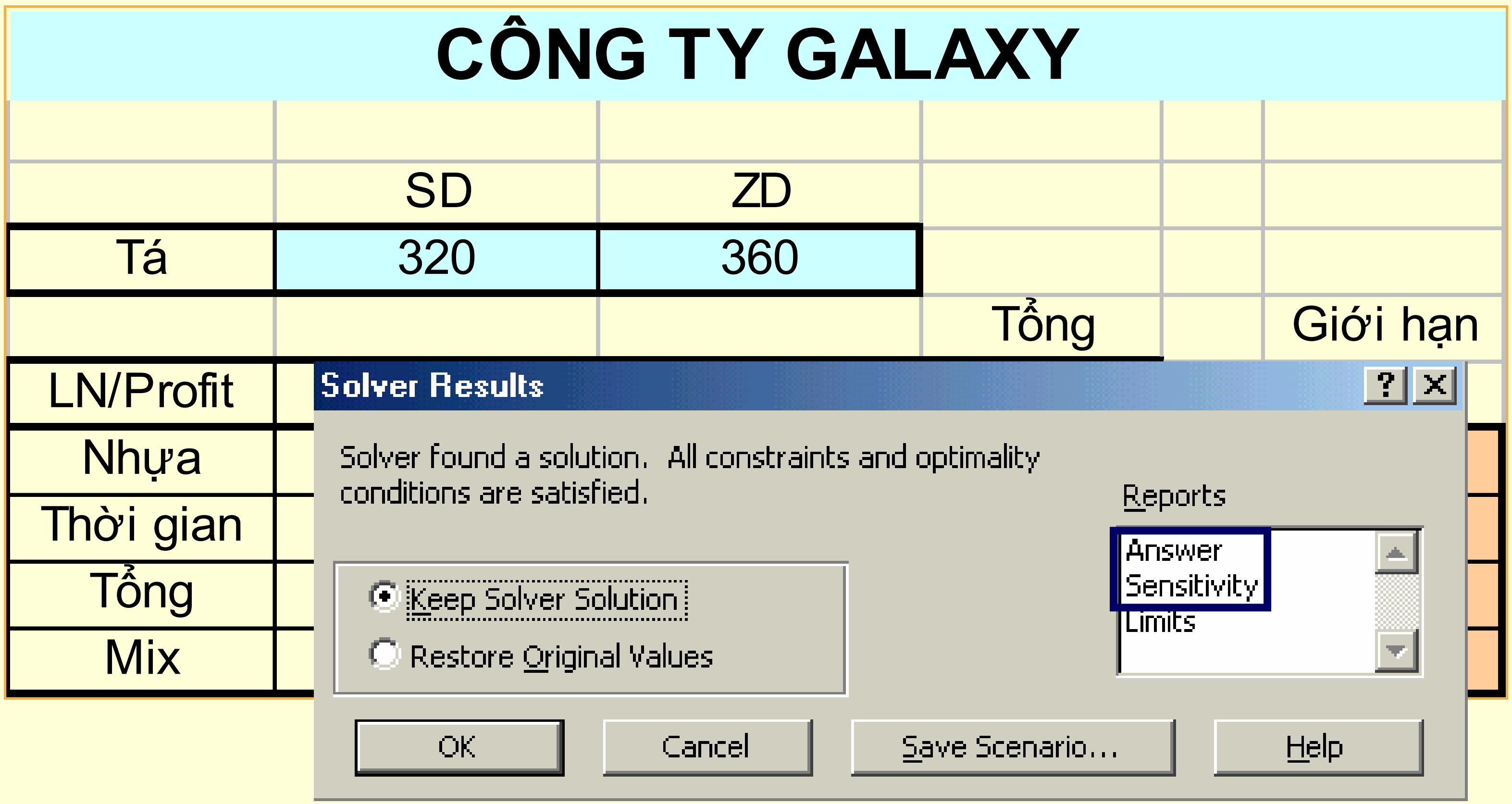
- Excel: Galaxy.xls

CÔNG TY GALAXY					
	SD	ZD			
Tá	320	360			
			Tổng		Giới hạn
LN/Profit	8	5	4360		
Nhựa	2	1	1000	<=	1000
Thời gian	3	4	2400	<=	2400
Tổng	1	1	680	<=	700
Mix	1	-1	-40	<=	350

Phân tích độ nhạy cho bài toán quy hoạch và ứng dụng Solver

Sử dụng Excel Solver tìm phương án tối ưu

- Excel: Galaxy.xls



5.3. Mô hình qui hoạch tuyến tính

Phân tích độ nhạy cho bài toán quy hoạch và ứng dụng Solver

Sử dụng Excel Solver tìm phương án tối ưu

- Excel Solver –Answer Report

Microsoft Excel 11.0 Answer Report
Worksheet: [Galaxy Alt.xls]Alt
Report Created: 7/28/2009 3:07:40 PM

Target Cell (Max)

Cell	Name	Original Value	Final Value
$D$6	LN/Profit Tổng	4360	4360

Adjustable Cells

Cell	Name	Original Value	Final Value
$B$4	Tá SD	320	320
$C$4	Tá ZD	360	360

Constraints

Cell	Name	Cell Value	Formula	Status	Slack
$D$7	Nhựa Tổng	1000	$D$7<=$F$7	Binding	0
$D$8	Thời gian Tổng	2400	$D$8<=$F$8	Binding	0
$D$9	Sản xuất Tổng	680	$D$9<=$F$9	Not Binding	20
$D$10	Mix Tổng	-40	$D$10<=$F$10	Not Binding	390

Phân tích độ nhạy cho bài toán quy hoạch và ứng dụng Solver

Sử dụng Excel Solver tìm phương án tối ưu

- Excel Solver –Sensitivity Report

Microsoft Excel 11.0 Sensitivity Report
Worksheet: [Galaxy Alt.xls]Alt
Report Created: 7/28/2009 3:07:40 PM

Adjustable Cells

Cell	Name	Final Value	Reduced Cost	Objective Coefficient	Allowable Increase	Allowable Decrease
$B$4	Tá SD	320	0	8	2	4.25
$C$4	Tá ZD	360	0	5	5.666666667	1

Constraints

Cell	Name	Final Value	Shadow Price	Constraint R.H. Side	Allowable Increase	Allowable Decrease
$D$7	Nhựa Tổng	1000	3.4	1000	100	400
$D$8	Thời gian Tổng	2400	0.4	2400	100	650
$D$9	Sản xuất Tổng	680	0	700	1E+30	20
$D$10	Mix Tổng	-40	0	350	1E+30	390

Chương 6

PHÂN TÍCH RA QUYẾT ĐỊNH

6.1. Giới thiệu

Khái niệm về ra quyết định

- Ra quyết định là một quá trình lựa chọn có ý thức giữa hai hoặc nhiều phương án để chọn ra một phương án và phương án này sẽ tạo ra được một kết quả mong muốn trong các điều kiện ràng buộc đã biết.
- Lưu ý rằng, nếu chỉ có một giải pháp để giải quyết vấn đề thì không phải là bài toán ra quyết định. Và cũng cần lưu ý rằng, phương án "Không làm gì cả" cũng là một phương án, đôi khi đó lại là phương án được chọn

6.1. Giới thiệu

Các loại ra quyết định

- Theo tính chất của vấn đề, có thể chia quyết định làm ba loại:
  - Ra quyết định trong điều kiện chắc chắn (certainty): Khi ra quyết định, đã biết chắc chắn trạng thái nào sẽ xảy ra , do đó sẽ dễ dàng và nhanh chóng ra quyết định.
  - Ra quyết định trong điều kiện rủi ro (risk): Khi ra quyết định đã biết được xác suất xảy ra của mỗi trạng thái.
  - Ra quyết định trong điều kiện không chắc chắn (uncertainty): Khi ra quyết định, không biết được xác suất xảy ra của mỗi trạng thái hoặc không biết được các dữ liệu liên quan đến các vấn đề cần giải quyết.

6.1. Giới thiệu

Các bước của quá trình ra quyết định

- Quá trình ra quyết định thường được tiến hành theo sáu bước:
  - Bước 1: Xác định rõ vấn đề cần giải quyết.
  - Bước 2: Liệt kê tất cả các phương án có thể có.
  - Bước 3: Nhận ra các tình huống hay các trạng thái.
  - Bước 4: Ước lượng tất cả lợi ích và chi phí cho mỗi phương án ứng với mỗi trạng thái.
  - Bước 5: Lựa chọn một mô hình toán học trong PP định lượng để tìm lời giải tối ưu.
  - Bước 6: Áp dụng mô hình để tìm lời giải và dựa vào đó để ra quyết định .

6.1. Giới thiệu

Ví dụ bài toán ra quyết định

- Ông A là Giám đốc của công ty X muốn ra quyết định về một vấn đề sản xuất, ông lần lượt thực hiện sáu bước như sau:
  - <u>Bước 1</u>: Ông A nêu vấn đề có nên sản xuất một sản phẩm mới để tham gia thị trường hay không?
  - <u>Bước 2</u>: Ông A cho rằng có 3 phương án sản xuất là:
    - *Phương án 1: lập 1 nhà máy có qui mô lớn để sản xuất sản phẩm.*
    - *Phương án 2: lập 1 nhà máy có qui mô nhỏ để sản xuất sản phẩm.*
    - *Phương án 3: không làm gì cả.*
  - <u>Bước 3</u>: Ông A cho rằng có 2 tình huống của thị trường sẽ xảy ra là:
    - *Thị trường tốt.*
    - *Thị trường xấu.*

Ví dụ bài toán ra quyết định

- Bước 4: Ông A ước lượng lợi nhuận của các phương án ứng với các tình huống:

Bảng 6.1

Phương án	Trạng thái	
	Thị trường Tốt	Thị trường Xấu
Nhà máy lớn	200.000	- 180.000
Nhà máy nhỏ	100.000	- 20.000
Không làm gì	0	0

- Bước 5 và 6: Chọn một mô hình toán học trong phương pháp định lượng để tác dụng vào bài toán này. Việc chọn lựa mô hình được dựa vào sự hiểu biết, vào thông tin ít hay nhiều về khả năng xuất hiện các trạng thái của hệ thống.

6.2. Ra quyết định trong điều kiện rủi ro

Khái niệm

- Khi ra quyết định trong điều kiện rủi ro, ta đã biết được xác suất xảy ra của mỗi trạng thái. Ra quyết định trong điều kiện rủi ro, ta thường sử dụng các tiêu chuẩn sau:
  - Cực đại giá trị kỳ vọng được tính bằng tiền EMV (Expected Moneytary Value), hay;
  - Cực tiểu thiệt hại kỳ vọng EOL (Expected Opportunity Loss).
- Để xác định các tiêu chuẩn trên người ta có thể sử dụng phương pháp lập bảng quyết định hoặc cây quyết định.

6.2. Ra quyết định trong điều kiện rủi ro

Phương pháp lập bảng quyết định

- ***Mô hình Max EMV(i)***:
  - Trong mô hình này, chúng ta sẽ chọn phương án i có giá trị kỳ vọng tính bằng tiền lớn nhất. EMV (i): giá trị kỳ vọng tính bằng tiền của phương án i.

$$EMV(i) = \sum_{j=1}^{m} P(S_j) x P_{ij}$$

  - Trong đó:
    - *$P(S_j)$: xác suất để trạng thái j xuất hiện.*
    - *P_{ij} : là lợi nhuận/chi phí của phương án i ứng với trạng thái j.*
    - *i = 1 đến n và j = 1 đến m.*

6.2. Ra quyết định trong điều kiện rủi ro

Phương pháp lập bảng quyết định

- ***Mô hình Max EMV(i)***:
  - Ví dụ: Trở lại bài toán của ông giám đốc A của công ty X với giả sử rằng thị trường xấu cũng như thị trường tốt đều có xác suất như nhau và bằng 0.5. Suy ra ta có bảng kết quả tương ứng:

Bảng 6.2

Phương án i	**Trạng thái j**		**EMV(i)**
	Thị trường tốt (j = 1)	**Thị trường xấu (j = 2)**	
Nhà máy lớn (i=1)	200.000	-180.000	10.000
Nhà máy nhỏ (i=2)	100.000	-20.000	40.000
Không làm gì (i=3)	0	0	0
Xác suất các trạng thái P(Sj)	0,5	0,5	

  - Ra quyết định:
    - *EMV (i) > 0 => phương án có lợi*
    - *Max EMV (i) =EMV (i=2) = 40.000 => Chọn phương án nhà máy nhỏ.*

6.2. Ra quyết định trong điều kiện rủi ro

Phương pháp lập bảng quyết định

- ***<u>Mô hình Min EOL(i) (Expeded Opportunity Loss, Thiệt hại cơ hội kỳ vọng)</u>***:
  - OL_{ij} là thiệt hại cơ hội của phương án i ứng với trạng thái j được định nghĩa như sau:

$$OL_{ij} = MaxP_{ij} - P_{ij}$$

  - Đây cũng chính là số tiền ta bị thiệt hại khi ta không chọn được phương án tối ưu mà phải chọn phương án i.
  - *<u>Ví dụ:</u>* Từ bảng 6.2 ta có:
    - $OL_{11} = 200.000 - 200.000 = 0$
    - $OL_{12} = 0 - (-180.000) = 180.000$
    - $OL_{21} = 200.000 - 100.000 = 100.000$
    - $OL_{22} = 0 - (-20.000) = 20.000$
    - $OL_{31} = 200.000 - 0 = 200.000$
    - $OL_{22} = 0 - 0 = 0$

6.2. Ra quyết định trong điều kiện rủi ro

Phương pháp lập bảng quyết định

- ***Mô hình Min EOL(i) (Expeded Opportunity Loss, Thiệt hại cơ hội kỳ vọng)***:
  - *Ví dụ:* Do đó ta có bảng trạng thái:

Bảng 6.3

Phương án	**Trạng thái**	
	Thị trường Tốt	**Thị trường Xấu**
Nhà máy lớn	0	180.000
Nhà máy nhỏ	100.000	20.000
Không làm gì	200.000	0
Xác suất của các trạng thái	0,5	0,5

  - Thiệt hại cơ hội kỳ vọng EOL(i) (Expected Opportunity loss):

$$EOL(i) = \sum_{j=1}^{m} P(S_j) x OL_{ij}$$

6.2. Ra quyết định trong điều kiện rủi ro

Phương pháp lập bảng quyết định

- ***Mô hình Min EOL(i) (Expeded Opportunity Loss, Thiệt hại cơ hội kỳ vọng)***:
  - *Ví dụ:* Từ bảng trạng thái, ta tính được:
    - *EOL (lớn) = 0.5 x 0 + 0 .5 x 180.000 = 90.000*
    - *EOL (nhỏ) = 0.5 x 100.000 + 0.5 x 20.000 = 60.000*
    - *EOL (không) = 0.5 x 200.000 + 0.5 x 0 = 100.000*
  - Ra quyết định theo tiêu chuẩn Min EOL (i)
  - Min EOL (i) = Min (90.000, 60.000, 100.000) = 60.000
  - => Chọn phương án nhà máy nhỏ

6.2. Ra quyết định trong điều kiện rủi ro

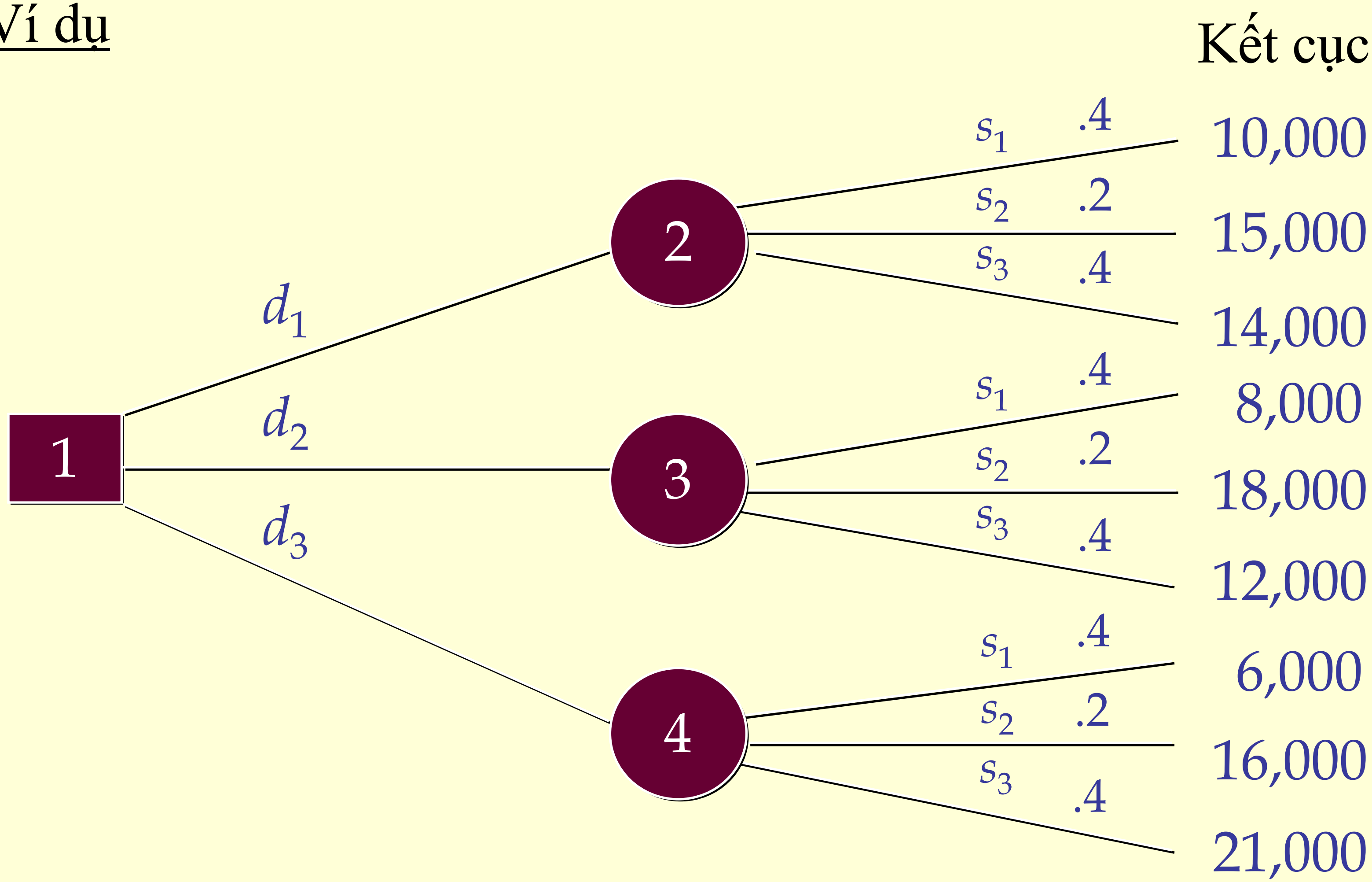
Phương pháp cây quyết định

- Có rất nhiều bài toán mà để đi đến quyết định cuối cùng phải qua một loạt các quyết định liên kết với nhau. Một phân tích giá trị mong đợi có thể dễ dàng được mở rộng để áp dụng cho những bài toán phức tạp dạng này, phương pháp này được gọi là "Phân tích cây quyết định".
- Logic và nguyên lý của phương pháp này vẫn giống như trong phân tích giá trị dự tính nhưng các bài toán thì phức tạp hơn.

6.2. Ra quyết định trong điều kiện rủi ro

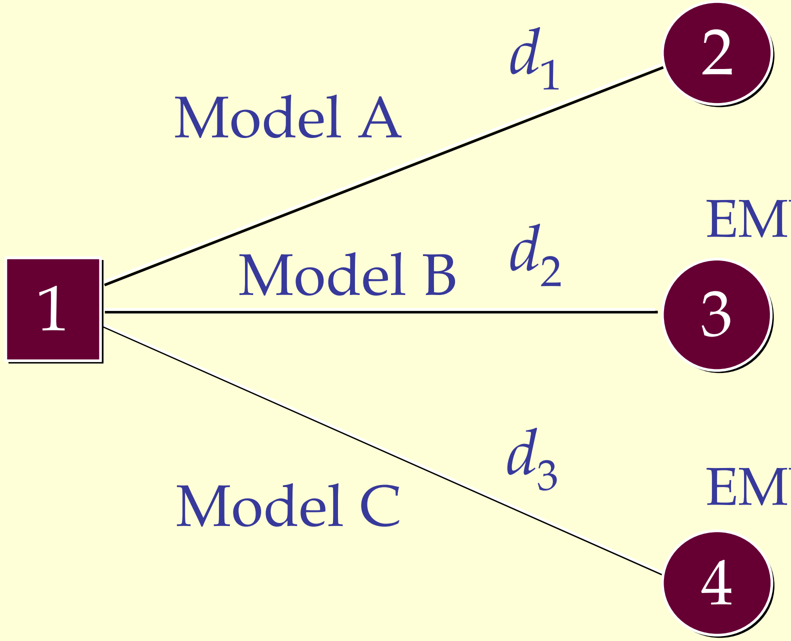
Phương pháp cây quyết định

Ví dụ



Phương pháp cây quyết định

<u>Ví dụ</u>



EMV = .4(10,000) + .2(15,000) + .4(14,000)
= $12,600

EMV = .4(8,000) + .2(18,000) + .4(12,000)
= $11,600

EMV = .4(6,000) + .2(16,000) + .4(21,000)
= $14,000

Lựa chọn mô hình có EMV lớn nhất là mô hình C

6.2. Ra quyết định trong điều kiện rủi ro

Phương pháp cây quyết định

Ví dụ:

	A	B	C	D	E	F
1	BẢNG PHÂN TÍCH					
2						
3	Mô hình	Các khả năng			EMV	Quyết định
5	d1 = Model A	10,000	15,000	14,000	=$B$8*B5+$C$8*C5+$D$8*D5	=IF(E5=$E$9,A5,"")
6	d2 = Model B	8,000	18,000	12,000	=$B$8*B6+$C$8*C6+$D$8*D6	=IF(E6=$E$9,A6,"")
7	d3 = Model C	6,000	16,000	21,000	=$B$8*B7+$C$8*C7+$D$8*D7	=IF(E7=$E$9,A7,"")
8	Xác suất	0.4	0.2	0.4		
9		Maximum Expected Value			=MAX(E5:E7)	

	A	B	C	D	E	F
1	BẢNG KẾT QuẢ					
2						
3	Mô hình	Các khả năng			EMV	Quyết định
5	d1 = Model A	10,000	15,000	14,000	12600	
6	d2 = Model B	8,000	18,000	12,000	11600	
7	d3 = Model C	6,000	16,000	21,000	14000	**d3 = Model C**
8	Xác suất	0.4	0.2	0.4		
9		Maximum Expected Value			**14000**	

6.3. Quyết định trong điều kiện không chắc chắn

- Trong điều kiện không chắc chắn, ta không biết được xác suất xuất hiện của mỗi trạng thái hoặc các dữ kiện liên quan đến bài toán không có sẵn. Trong trường hợp này ta có thể dùng một trong 5 mô hình sau:
  - Maximax
  - Maximin
  - Đồng đều ngẫu nhiên (Equally -likely)
  - Tiêu chuẩn hiện thực hay tiêu chuẩn Hurwiez
  - Minimax

- *Ghi chú:*
  - *Bốn mô hình đầu được tính từ bảng 6.1*
  - *Mô hình cuối cùng được tính từ bảng 6.3*

6.3. Quyết định trong điều kiện không chắc chắn

Mô hình Maximax

- Tìm phương án i ứng với Max của max, nghĩa là tìm giá trị lớn nhất trong bảng quyết định

$$\underset{i}{Max}(\underset{j}{Max}\,P_{ij})$$

- Trong mô hình này ta tìm lợi nhuận tối đa có thể có được bất chấp rủi ro, vì vậy tiêu chuẩn này còn được gọi là tiêu chuẩn lạc quan (optimistic decision criterion).
- Ví dụ:
  - Từ bảng 6.1 ta có $\underset{i}{Max}(\underset{j}{Max}\,P_{ij}) = 200.000$
  - Ra quyết định: chọn phương án nhà máy lớn

6.3. Quyết định trong điều kiện không chắc chắn

Mô hình Maximin

- Chọn phương án i ứng với Max của Min

$$\underset{i}{Max}(\underset{j}{Min}\,P_{ij})$$

- Nghĩa là tìm Min trong hàng i, sau đó lấy Max những giá trị Min vừa tìm được. Cách làm này phản ánh tinh thần bi quan, còn gọi là quyết định bi quan (pessimistic decision).
- <u>Ví dụ</u>:
  - Từ bảng 6.1 ta có $\underset{i}{Max}(\underset{j}{Min}\,P_{ij}) = 0$
  - Ra quyết định: không làm gì cả

6.3. Quyết định trong điều kiện không chắc chắn

Mô hình Đồng đều ngẫu nhiên

- Trong mô hình này, ta xem mọi trạng thái đều đồng đều ngẫu nhiên, nghĩa là xem các trạng thái đều có xác suất xuất hiện bằng nhau. Trong trường hợp này ta tìm phương án i ứng với:

$$\underset{i}{Max}\left(\frac{\sum_{j=1}^{m} P_{ij}}{So\ trang\ thai}\right)$$

- Nghĩa là tìm phương án làm cực đại giá trị trung bình các lợi nhuận của từng phương án.
- <u>Ví dụ</u>:
  - Từ bảng 6.1 ta có: $\underset{i}{Max}\left(\frac{200.000+(-180.000)}{2};\frac{100.000+(-20.000)}{2};\frac{0+0)}{2}\right)$

    $\underset{i}{Max}(10.000;40.000;0)$
  - Ra quyết định: Chọn phương án xây nhà máy nhỏ

6.3. Quyết định trong điều kiện không chắc chắn

Mô hình Hurwitz - còn được gọi là mô hình trung bình có trọng số

- Đây là mô hình dung hòa giữa tiêu chuẩn lạc quan và tiêu chuẩn bi quan. Bằng cách chọn một hệ số α (0< α <1). Sau đó chọn phương án i ứng với hệ số a sao cho:

$$\underset{i}{Max}[\alpha \underset{j}{Max}\, P_{ij} + (1-\alpha)\underset{j}{Min}\, P_{ij}]$$

  - $Min\ P_{ij}$: giá trị nhỏ nhất ở hàng thứ i
  - $Max\ P_{ij}$: giá trị lớn nhất ở hàng thứ i
  - Hệ số α : 0 < α <1
    - *α = 1: Người quyết định lạc quan về tương lai*
    - *α = 0: Người quyết định bi quan về tương lai*
- Phương pháp này có dạng mềm dẻo hơn, giúp cho người ra quyết định đưa được cảm xúc cá nhân về thị trường vào mô hình.

6.3. Quyết định trong điều kiện không chắc chắn

Mô hình Minimax

- Ta tìm phương án ứng với:

$$\underset{i}{Min}(\underset{j}{Max}\, OL_{ij})$$

  - *Tìm Max theo phương án i nghĩa là tìm giá trị lớn nhất trong các cột j tính theo từng hàng*
  - *OL_{ij}: thiệt hại cơ hội của phương án i ứng với trạng thái j được tính như trong mô hình ra quyết định trong điều kiện rủi ro.*
  - *Trong mô hình này ta tìm phương án để làm cực tiểu cơ hội thiệt hại cực đại.*
- *Ví dụ :* Áp dụng bảng 6.3 ta có:
  - Min [Max OL_{ij}]= Min [180.000 , 100.000 , 200.000]= 100.000
  - Ra quyết định: Chọn phương án nhà máy có qui mô nhỏ.

6.4. Quyết định khi xét đến độ hữu ích (độ thoả dụng)

Khái niệm độ hữu ích

Ví dụ:

- *Chọn 1 trong 2 phương án đầu tư sau:*
  - A : Chắc chắn thu được $30000
  - B : 70% khả năng thu được $60000, 30% khả năng thua lỗ $10000 (- $10000).

?

  - Dùng tiêu chuẩn EMV để đánh giá lựa chọn?
  - Thực tế lựa chọn?

6.4. Quyết định khi xét đến độ hữu ích (độ thoả dụng)

Khái niệm độ hữu ích

- Độ hữu ích là độ đo mức ưu tiên của người ra quyết định đối với lợi nhuận.
- Lý thuyết độ hữu ích là lý thuyết nghiên cứu cách kết hợp mức độ ưu tiên về độ may rủi của người ra quyết định đối với các yếu tố khác trong quá trình ra quyết định.
- ***<u>Độ hữu ích được ước tính như sau:</u>***
  - Kết quả tốt nhất sẽ có độ hữu ích là 1 => U (tốt nhất) = 1
  - Kết quả xấu nhất sẽ có độ hữu ích là 0 => U (xấu nhất) = 0
  - Kết quả khác sẽ có độ hữu ích $\in (0,1)$ => $0 <$ U(khác) < 1

6.4. Quyết định khi xét đến độ hữu ích (độ thoả dụng)

Khái niệm độ hữu ích

- *Vấn đề đặt ra 1*: Lựa chọn phương án nào?
  - A : Chắc chắn thu được $30000
  - B : 70% khả năng thu được $60000, 30% khả năng thua lỗ $10000 (- $10000).
  - *Khả năng lựa chọn*: 30% khả năng lỗ $10 000 là quá cao, quyết định chọn phương án chắc chắn.
- *Vấn đề đặt ra 2*: Lựa chọn phương án nào?
  - A : Chắc chắn thu được $30000
  - B : 90% khả năng thu được $60000, 10% khả năng thua lỗ $10000 (- $10000).
  - *Khả năng lựa chọn*: Cơ hội để có được $60000 là khá cao, quyết định chọn phương án được $60000.
- *Vấn đề đặt ra 3*: Lựa chọn phương án nào?
  - A : Chắc chắn thu được $30000
  - B : 85% khả năng thu được $60000, 15% khả năng thua lỗ $10000 (- $10000).
  - *Khả năng lựa chọn*: 2 phương án này thấy rằng tương đương nhau, chọn phương án nào cũng được.

6.4. Quyết định khi xét đến độ hữu ích (độ thoả dụng)

Khái niệm độ hữu ích

- Khi đó ta thấy rằng:

  30000 ~ {(0.85, 60000), (1-0.85, -10000) }

  ~ {(0.85, 60000), (0.15, -10000) }

- Độ hữu ích của 30000 được xác định như sau:
  - $u(\$30\,000) = 0.85u(\$60\,000) + 0.15u(-\$10\,000) = 0.85.$
  - *Ghi chú*: điểm 85% được gọi là điểm không khác biệt.

6.4. Quyết định khi xét đến độ hữu ích (độ thoả dụng)

Khái niệm độ hữu ích

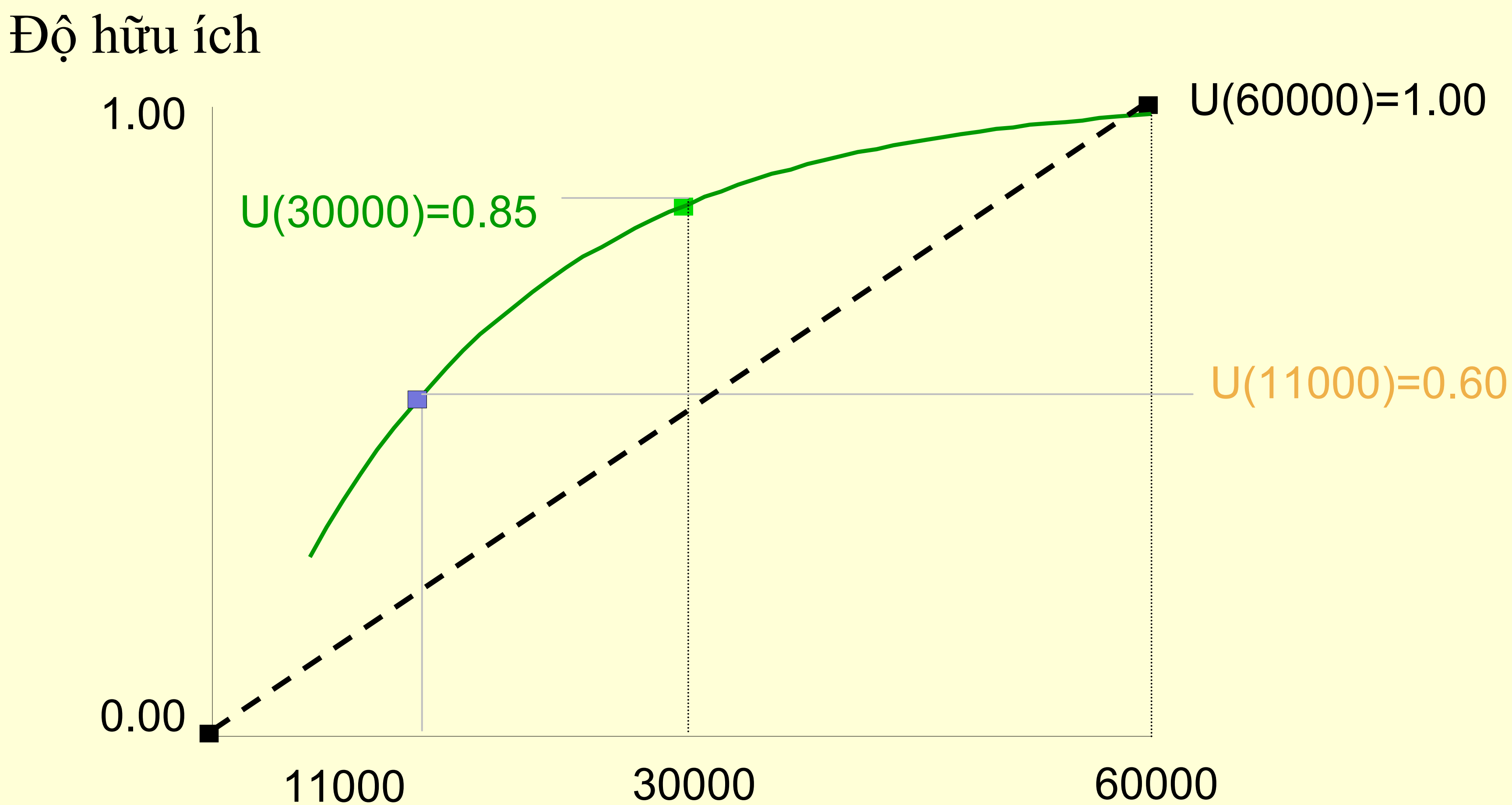
- *Tương tự, giả sử:*
- *Vấn đề đặt ra*: Lựa chọn phương án nào?
  - A : Chắc chắn thu được \$11000
  - B : 60% khả năng thu được \$60000, 40% khả năng thua lỗ \$10000 (- \$10000).
  - *Khả năng lựa chọn*: 2 phương án này thấy rằng tương đương nhau, chọn phương án nào cũng được.
  - Khi đó:

$$11000 \sim \{(0.6, 60000), (1-0.6, -10000)\}$$
$$\sim \{(0.6, 60000), (0.4, -10000)\}$$

- *Do đó:* $u(\$11\,000) = 0.6u(\$60\,000) + 0.4u(-\$10\,000) = 0.6.$

6.4. Quyết định khi xét đến độ hữu ích (độ thoả dụng)

Khái niệm độ hữu ích



6.4. Quyết định khi xét đến độ hữu ích (độ thoả dụng)

Khái niệm độ hữu ích

Đường độ hữu ích/ Utility curve: Risk Neutral, Averse, Seeking

Thank You !