

Chương VI

DÃY SỐ THỜI GIAN

Cấu trúc chương

1. Khái niệm và phân loại
2. Các chỉ tiêu phân tích dãy số thời gian
3. Một số phương pháp biểu diễn xu hướng cơ bản của hiện tượng
4. Một số phương pháp dự đoán thống kê ngắn hạn

1. Khái niệm và phân loại

a. Khái niệm

- *Là dãy các trị số của chỉ tiêu thống kê được sắp xếp theo thứ tự thời gian*
- *Ví dụ*

Năm	1997	1998	1999	2000	2001	2002
Giá trị XK (triệu USD)	10,0	10,2	11,0	11,8	13,0	14,8

**Ví dụ: Có tài liệu về lao động
của doanh nghiệp A như sau:**

Ngày	1.1.09	1.2.09	1.3.09	1.4.09
Số lao động (Người)	350	370	370	380

b. Cấu tạo - 2 thành phần sau:

- Thời gian: có thể là ngày, tuần, tháng, quý, năm tùy thuộc vào đặc điểm, tính chất của hiện tượng nghiên cứu. Độ dài giữa 2 thời gian liên nhau gọi là khoảng cách thời gian.
- Chỉ tiêu về hiện tượng nghiên cứu: tên, đơn vị tính phù hợp và trị số của chỉ tiêu. Các trị số này được gọi là các mức độ của dãy số thời gian y_i . Các mức độ của dãy số thời gian có thể là số tuyệt đối, số tương đối hoặc số bình quân.

Yêu cầu khi xây dựng DSTG:

- Phải thống nhất về nội dung và phương pháp tính chỉ tiêu qua thời gian.
 - Ví dụ: Chỉ tiêu GDP ở nước ta hiện nay tính theo Hệ thống tài khoản quốc gia của Liên hợp quốc, trước đó là chỉ tiêu tính theo Hệ thống sản xuất vật chất của Liên Xô cũ (MPS).
- Phải thống nhất về phạm vi tổng thể nghiên cứu.
 - Ví dụ: Từ 1/8/2008, Hà Nội bao gồm Hà Tây và một số địa phương thuộc Vĩnh Phúc, Hoà Bình. Không thể đem các số liệu của Hà Nội trước khi nhập tỉnh để so sánh với số liệu của Hà Nội hiện nay được.
- Các khoảng cách thời gian trong dãy số nên bằng nhau, nhất là với các dãy số thời kỳ phải bằng nhau.

Phân loại DSTG

Dãy số thời kỳ

- Là dãy số mà mỗi mức độ của nó biểu hiện quy mô, khối lượng của hiện tượng trong từng khoảng thời gian nhất định

Đặc điểm:

- Khoảng cách thời gian ảnh hưởng đến mức độ
- Có thể cộng dồn các mức độ

Dãy số thời điểm

- Là dãy số mà mỗi mức độ của nó biểu hiện quy mô, khối lượng của hiện tượng tại một thời điểm nhất định.

Đặc điểm

- Mức độ phản ánh quy mô tại thời điểm
- Không thể cộng dồn các mức độ

Ví dụ

Năm	1997	1998	1999	2000	2001	2002
Giá trị XK (triệu USD)	10,0	10,2	11,0	11,8	13,0	14,8

Ngày	1/4/03	1/5/03	1/6/03	1/7/03
GT tồn kho (tr\$)	3560	3640	3700	3540

Căn cứ vào các loại chỉ tiêu

- **Dãy số số tuyệt đối:** dãy số có các trị số của chỉ tiêu là số tuyệt đối.
 - Ví dụ: Qui mô vốn của doanh nghiệp qua các năm.
- **Dãy số số tương đối:** dãy số mà các trị số là các số tương đối.
 - Ví dụ: tốc độ tăng trưởng kinh tế hàng năm.
- **Dãy số số bình quân:** dãy số mà các trị số là các số bình quân.
 - Ví dụ: NSLĐ trung bình của doanh nghiệp qua các năm.

Ý nghĩa của việc nghiên cứu dãy số thời gian

- Nghiên cứu các đặc điểm về sự biến động của hiện tượng qua thời gian
- Phát hiện xu hướng phát triển và tính quy luật của hiện tượng
- Dự đoán mức độ của hiện tượng trong tương lai

2. Các chỉ tiêu phân tích DSTG

- a) Mức độ bình quân theo thời gian
- b) Lượng tăng/giảm tuyệt đối
- c) Tốc độ phát triển
- d) Tốc độ tăng/giảm
- e) Giá trị tuyệt đối của 1% tăng/giảm liên hoàn

a. Mức độ bình quân theo thời gian

Mức độ bình quân đối với DS thời kỳ

- Sử dụng số bình quân cộng giản đơn
 - Công thức:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

Ví dụ

Năm	1997	1998	1999	2000	2001	2002
Giá trị XK (triệu USD)	10,0	10,2	11,0	11,8	13,0	14,8
GTXK bình quân (tr. \$)	$(10,0+10,2+11,0+11,8+13,0+14,8)/6$ 11,8					

Mức độ bình quân theo thời gian

Mức độ bình quân đối với DS thời điểm

■ *Điều kiện để có thể tính được mức độ bình quân:*

- Mức độ cuối cùng của khoảng cách thời gian trước bằng mức độ đầu tiên của khoảng cách thời gian sau
- Giữa các thời điểm ghi chép số liệu, hiện tượng biến động tương đối đều đặn

■ **Phương pháp tính (k/c thời gian bằng nhau)**

- Tính mức độ bình quân của từng khoảng cách thời gian: bằng số bình quân của từng nhóm 2 mức độ
- Xác định mức độ bình quân trong cả giai đoạn: bằng số bình quân của các mức độ bình quân từng khoảng cách
- Bản chất của cách tính này là chuyển từ dãy số thời điểm sang dãy số thời kỳ để thực hiện phép tính

Ví dụ:

Ngày	1/4/03	1/5/03	1/6/03	1/7/03
GT hàng tồn kho (tr\$)	3560	3640	3700	3540

Xác định mức độ bình quân trong từng khoảng thời gian

Ngày	1/4/03	1/5/03	1/6/03	1/7/03
GT tồn kho (\$)	3560	3640	3700	3540
Mức độ bình quân từng khoảng cách (\$)	3600	3670	3620	

- GT hàng tồn kho bình quân trong Quý II/03 là mức độ bình quân của các tháng trong quý:
- GTTK bình quân : $(3600+3670+3620)/3 = \mathbf{3630}$ (\$)

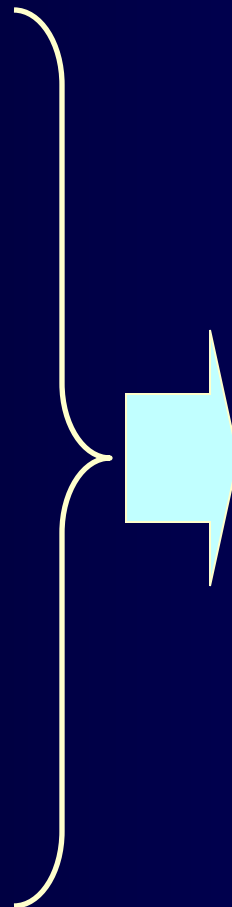
Công thức tổng quát

$$\bar{X}_{k/c1} = \frac{x_1 + x_2}{2}$$

$$\bar{X}_{k/c2} = \frac{x_2 + x_3}{2}$$

$$\bar{X}_{k/c3} = \frac{x_3 + x_4}{2}$$

$$\bar{X}_{n-1} = \frac{x_{n-1} + x_n}{2}$$



$$\bar{X} = \frac{\frac{x_1}{2} + x_2 + x_3 + \dots + \frac{x_n}{2}}{n-1}$$

Phương pháp tính: (k/c thời gian không bằng nhau)

Ví dụ:

- Thống kê tình hình nhân lực tại CT X tháng 4/03:
- Ngày 1 tháng 4 xí nghiệp có 400 công nhân
- Ngày 10 tháng 4 bổ sung 5 công nhân
- Ngày 16 tháng 4 bổ sung thêm 3 công nhân
- Ngày 21 tháng 4 cho 6 công nhân thôi việc, từ đó đến cuối tháng 4 không có gì thay đổi.

Phương pháp tính

(k/c thời gian không bằng nhau)

	Số ngày (f_i)	Số lượng CN (x_i)	$x_i f_i$
Từ 1 đến 9/4	9	400	3600
Từ 10 đến 15/4	6	405	2430
Từ 16 đến 20/4	5	408	2040
Từ 21 đến 30/4	10	402	4020
Tổng	30	x	12090

Số lượng công nhân bq tháng 4/03: $12090/30 = 403$ (CN)

Công thức tổng quát

$$\bar{x} = \frac{\sum_{i=1}^n \bar{x}_i f_i}{\sum_{i=1}^n f_i}$$

Trong đó:

- $\bar{x}_i$: mức độ bình quân của k/c thời gian i
- f_i : độ dài tương đối của k/c thời gian i
- n: số khoảng cách thời gian được theo dõi

b. Lượng tăng/giảm tuyệt đối (δ):

Lượng tăng/giảm tuyệt đối liên hoàn (δ_i)

- KN: Là chênh lệch giữa mức độ của kỳ nghiên cứu so với mức độ của kỳ đứng liền trước đó
- δ_i cho biết lượng tăng/giảm bằng số tuyệt đối của hiện tượng giữa hai kỳ quan sát liền nhau
- Công thức: $\delta_i = x_i - x_{i-1} (i=\overline{2,n})$

Ví dụ

Năm	1997	1998	1999	2000	2001	2002
Giá trị XK (triệu USD)	10,0	10,2	11,0	11,8	13,0	14,8
δ_i (tr\$)	-	0,2	0,8	0,8	1,2	1,8

Lượng tăng/giảm tuyệt đối định gốc Δ_i

- *Là chênh lệch giữa mức độ kỳ nghiên cứu với mức độ kỳ được chọn làm gốc cố định.*
- Δ_i cho thấy lượng tăng/giảm bằng số tuyệt đối của hiện tượng giữa kỳ nghiên cứu với gốc so sánh
- CT: $\Delta_i = x_i - x_1 (i=\overline{2,n})$
- Thực tế thì có thể chọn bất kỳ gốc nào nhưng về mặt lý thuyết thì thường chọn gốc đầu tiên

Nhận xét quan hệ giữa các δ_i và Δ_n

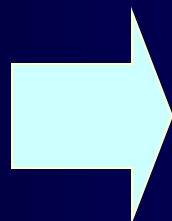
$$\blacksquare \delta_2 = x_2 - x_1$$

$$\blacksquare \delta_3 = x_3 - x_2$$

$$\blacksquare \delta_4 = x_4 - x_3$$

$$\blacksquare \dots\dots\dots$$

$$\blacksquare \delta_n = x_n - x_{n-1}$$



$$\Sigma \delta_i = x_n - x_1 = \Delta_n$$

→ Lượng tăng/giảm tuyệt đối định gốc kỳ nghiên cứu bằng tổng các lượng tăng/giảm tuyệt đối liên hoàn tính tới kỳ nghiên cứu

Ví dụ

Năm	1997	1998	1999	2000	2001	2002
Giá trị XK (triệu USD)	10,0	10,2	11,0	11,8	13,0	14,8
δ_i (tr\$)	-	0,2	0,8	0,8	1,2	1,8
Δ_i (tr\$)	-	0,2	1,0	1,8	3,0	4,8

Lượng tăng/giảm tuyệt đối bình quân $\bar{\delta}$

- Là số bình quân của các lượng tăng/giảm tuyệt đối liên hoàn
- $\bar{\delta}$ cho thấy mức độ đại diện về lượng tăng/giảm tuyệt đối qua các kỳ
- Công thức:
- $\bar{\delta}$ chỉ phụ thuộc vào mức độ đầu tiên và mức độ cuối cùng. Do vậy chỉ nên tính khi các mức độ của dãy số có cùng xu hướng và nên kết hợp với các lượng tăng giảm tuyệt đối liên hoàn δ_i để phân tích thì mới chặt chẽ

$$\bar{\delta} = \frac{\sum_{i=2}^n \delta_i}{n-1}$$

$$\bar{\delta} = \frac{\Delta_n}{n-1}$$

Ví dụ

Năm	1997	1998	1999	2000	2001	2002
Giá trị XK (triệu USD)	10,0	10,2	11,0	11,8	13,0	14,8
δ_i (tr\$)	-	0,2	0,8	0,8	1,2	1,8
Δ_i (tr\$)	-	0,2	1,0	1,8	3,0	4,8
$\bar{\delta}$ (tr\$)	0,96					

c. Tốc độ phát triển (t):

Tốc độ phát triển liên hoàn là tỷ số giữa mức độ kỳ nghiên cứu với mức độ kỳ liên trước đó.

- Chỉ tiêu này phản ánh sự phát triển của hiện tượng giữa hai thời gian liên nhau
- CT: $t_i = x_i / x_{i-1}$ ($i=2, \overline{n}$)
- Đơn vị: (lần) hoặc (%)

Ví dụ

N ^m	1997	1998	1999	2000	2001	2002
x _i (tr\$)	10,0	10,2	11,0	11,8	13,0	14,8
t _i (%)	-	102,0	107,8	107,3	110,2	113,8

Tốc độ phát triển định gốc (T_i)

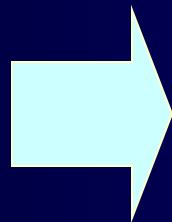
- *Là tỷ số giữa mức độ kỳ nghiên cứu với mức độ kỳ được chọn làm gốc.*
- T_i phản ánh sự phát triển của hiện tượng so với kì được chọn làm gốc đối chiếu
- Công thức tính: $T_i = x_i/x_1$ ($i = \overline{2,n}$)

Ví dụ

Năm	1997	1998	1999	2000	2001	2002
x_i (tr\$)	10,0	10,2	11,0	11,8	13,0	14,8
t_i (%)	-	102,0	107,8	107,3	110,2	113,8
T_i (%)	-	102,0	110,0	118,0	130,0	148,0

Nhận xét quan hệ giữa các t_i và T_n

- $t_2 = x_2/x_1$
- $t_3 = x_3/x_2$
- $t_4 = x_4/x_3$
-
- $t_n = x_n/x_{n-1}$



$$\prod t_i = x_n/x_1 = T_n$$

- $\rightarrow$ Tốc độ phát triển định gốc kỳ nghiên cứu bằng tích các tốc độ phát triển liên hoàn tính tới kỳ nghiên cứu

Tốc độ phát triển bình quân ($\bar{t}$)

- KN: Là số bình quân của các tốc độ phát triển liên hoàn
- Tốc độ phát triển bình quân cho thấy mức độ đại diện của tốc độ phát triển trong khoảng thời gian đó
- Chỉ nên tính khi các mức độ của dãy số có cùng xu hướng. Nếu không cùng xu hướng thì nên dùng tốc độ phát triển liên hoàn
- CT

$$\bar{t} = \sqrt[n-1]{\prod_{i=2}^n t_i}$$

$$\bar{t} = \sqrt[n-1]{T_n}$$

$$\bar{t} = \sqrt[n-1]{\frac{x_n}{x_1}}$$

Ví dụ

Năm	1997	1998	1999	2000	2001	2002
x_i (tr\$)	10,0	10,2	11,0	11,8	13,0	14,8
t_i (%)	-	102,0	107,8	107,3	110,2	113,8
$\bar{t}$ (%)	108,16					

d. Tốc độ tăng/giảm

Tốc độ tăng/giảm liên hoàn (a_i)

- KN: là tỷ số so sánh giữa lượng tăng/giảm tuyệt đối liên hoàn với mức độ kỳ gốc liên hoàn
- a_i cho biết lượng tăng/giảm bằng *số tương đối của hiện tượng giữa hai kỳ quan sát liên nhau*
- CT

$$a_i = \frac{\delta_i}{x_{i-1}}$$

$$a_i = \frac{x_i - x_{i-1}}{x_{i-1}}$$

$$a_i = t_i - 1(100\%)$$

Ví dụ

Năm	1997	1998	1999	2000	2001	2002
x_i (tr\$)	10,0	10,2	11,0	11,8	13,0	14,8
t_i (%)	-	102,0	107,8	107,3	110,2	113,8
a_i (%)	-	2,0	7,8	7,3	10,2	13,8

Tốc độ tăng/giảm định gốc (A_i)

- KN: *Tốc độ tăng/giảm định gốc là tỷ số so sánh giữa lượng tăng/giảm định gốc với mức độ kỳ gốc cố định*
- A_i cho biết lượng tăng/giảm bằng *số tương đối của hiện tượng giữa kỳ nghiên cứu với kỳ gốc cố định*
- CT: $A_i = \Delta_i / y_1 = (y_i - y_1) / y_1 = T_i - 1$ (lần)
- Nếu T_i tính bằng % thì $A_i = T_i - 100$

Ví dụ

Năm	1997	1998	1999	2000	2001	2002
x_i (tr\$)	10,0	10,2	11,0	11,8	13,0	14,8
Δ_i (tr\$)	-	0,2	1,0	1,8	3,0	4,8
T_i (%)	-	102,0	110,0	118,0	130,0	148,0
A_i (%)	-	2,0	10,0	18,0	30,0	48,0

Tốc độ tăng/giảm bình quân

- KN:
- Là chỉ tiêu tương đối nói lên nhịp điệu tăng/giảm đại diện tại thời kỳ nhất định
- CT: $\bar{a} = \bar{t} - 1$ (100%)
- Cũng chỉ nên sử dụng tốc độ tăng/giảm bình quân khi dãy số có cùng xu hướng.

Ví dụ

Năm	1997	1998	1999	2000	2001	2002
x_i (tr\$)	10,0	10,2	11,0	11,8	13,0	14,8
$\bar{t}$ (%)	108,16					
$\bar{a}$ (%)	8,16					

e. Giá trị tuyệt đối của 1% tăng/giảm: g_i

- KN: 1% tăng hoặc giảm của tốc độ tăng/giảm liên hoàn thì tương ứng với trị số tuyệt đối là bao nhiêu?
- g_i là số tuyệt đối, nên đơn vị tính tương ứng theo đơn vị tính của chỉ tiêu nghiên cứu.
- CT

$$g_i = \frac{\delta_i}{a_i}$$

$$g_i = \frac{x_i - x_{i-1}}{\frac{x_i - x_{i-1}}{x_{i-1}} \times 100}$$

$$g_i = \frac{x_{i-1}}{100}$$

Ví dụ

Năm	1997	1998	1999	2000	2001	2002
x_i (tr\$)	10,0	10,2	11,0	11,8	13,0	14,8
δ_i (tr\$)	-	0,2	0,8	0,8	1,2	1,8
a_i (%)	-	2,0	7,8	7,3	10,2	13,8
g_i (tr\$)	-	0,100	0,102	0,110	0,118	0,130

- Trên thực tế, thường không dùng chỉ tiêu giá trị tuyệt đối của 1% tốc độ tăng (giảm) định gốc vì nó luôn là 1 hằng số

Bảng chỉ tiêu phân tích DSTG

Năm	1997	1998	1999	2000	2001	2002
x_i (\$)	10,0	10,2	11,0	11,8	13,0	14,8
$\bar{x}$ (\$)	11,8					
δ_i (\$)	-	0,2	0,8	0,8	1,2	1,8
Δ_i (\$)	-	0,2	1,0	1,8	3,0	4,8
$\bar{\delta}$ (\$)	0,96					
t_i (%)	-	102,0	107,8	107,3	110,2	113,8
T_i (%)	-	102,0	110,0	118,0	130,0	148,0
$\bar{t}$ (%)	108,16					
a_i (%)	-	2,0	7,8	7,3	10,2	13,8
A_i (%)	-	2,0	10,0	18,0	30,0	48,0
$\bar{a}$ (%)	8,16					
g_i (\$)	-	0,100	0,102	0,110	0,118	0,130

Bài tập 2 trang 354

3. Một số phương pháp biểu hiện xu hướng phát triển của hiện tượng

Mục đích chung của các phương pháp

- Hiện tượng biến động qua thời gian, chịu ảnh hưởng bởi nhiều nhóm nhân tố, trong đó:
 - Các nhân tố chủ yếu, tác động đến hiện tượng và quyết định xu hướng phát triển cơ bản của hiện tượng.
 - Các nhân tố ngẫu nhiên tác động một cách ngẫu nhiên làm cho hiện tượng sai lệch so với xu hướng chung.
- Vấn đề đặt ra là phải loại trừ những nhân tố ngẫu nhiên và làm bộc lộ ra những nhân tố cơ bản để phản ánh xu hướng phát triển của hiện tượng

1 – Phương pháp mở rộng khoảng cách thời gian

- Phạm vi áp dụng:

Dãy số thời gian có khoảng cách thời gian tương đối ngắn và có nhiều mức độ mà chưa biểu hiện được xu hướng phát triển của hiện tượng.

VD:

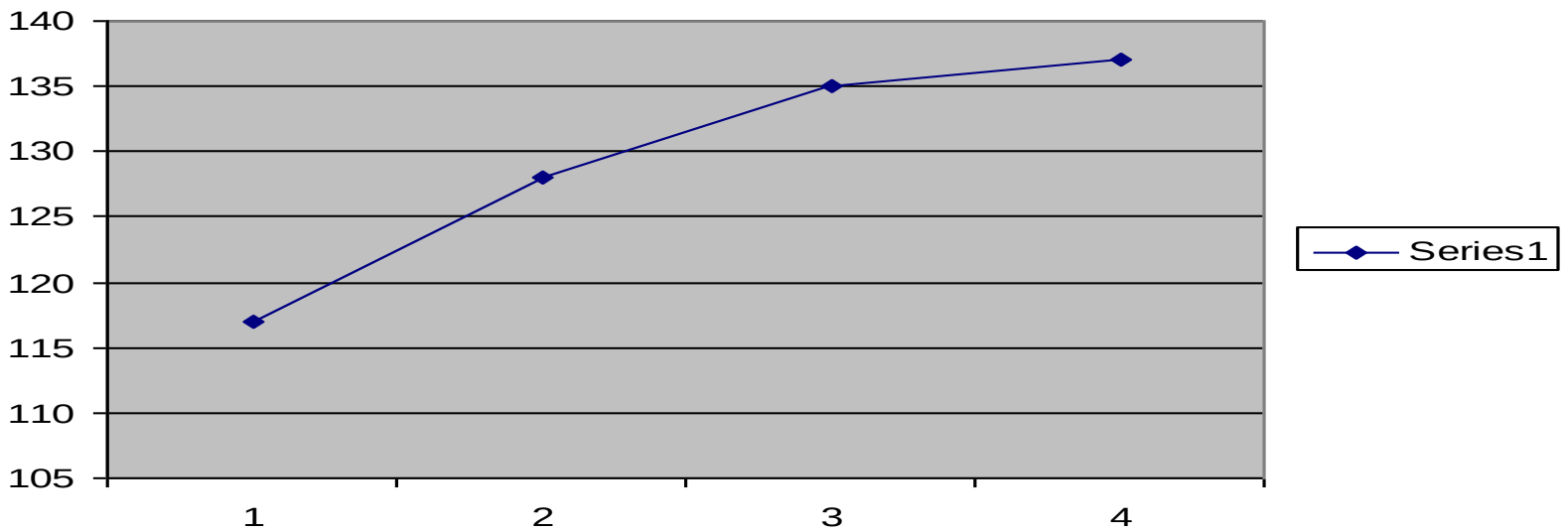
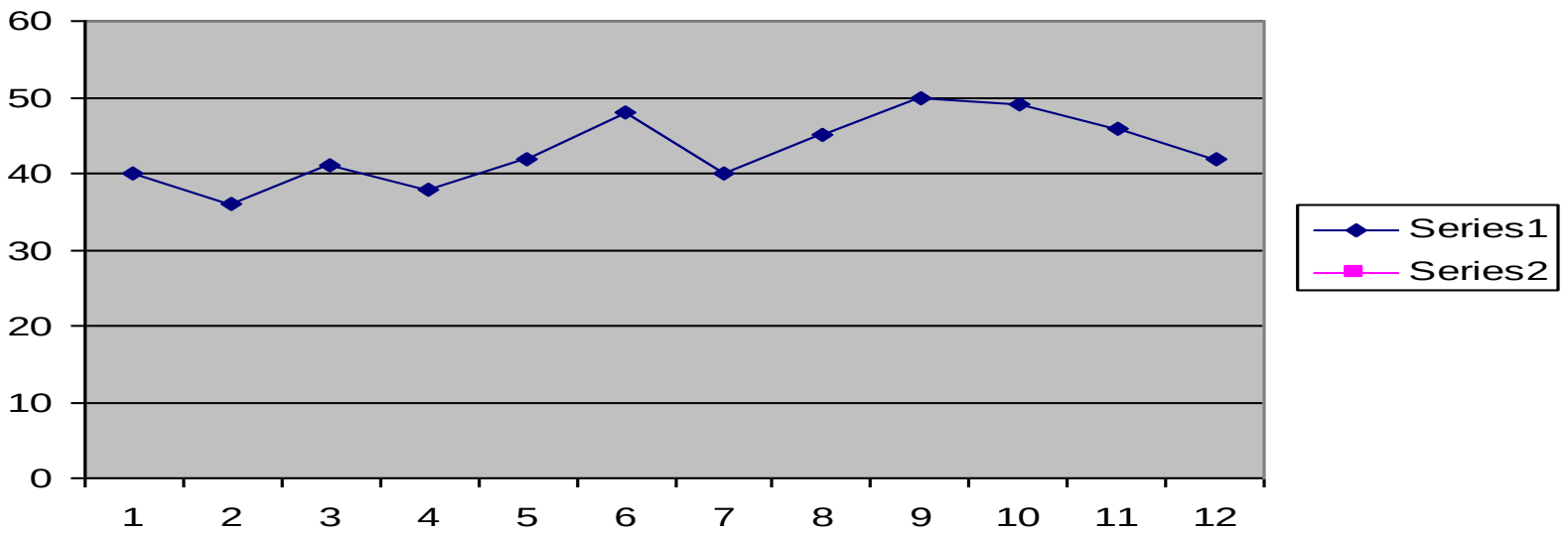
Tháng	1	2	3	4	5	6	7	8	9	10	11	12
Sản lượng (1000 tấn)	40	36	41	38	42	48	40	45	50	49	46	42

- **Nội dung của phương pháp**

- Từ dãy số thời gian đã cho xây dựng một dãy số thời gian mới bằng cách mở rộng thêm khoảng cách thời gian thông qua việc ghép một số thời gian liên nhau vào thành một khoảng thời gian dài hơn.
- Từ VD trên có thể mở rộng khoảng cách thời gian từ tháng sang quý để nhìn rõ xu hướng thay đổi của sản lượng.

Quý	I	II	III	IV
Sản lượng (1000 tấn)	117	128	135	137

- Thời gian dài - ngắn mang ý nghĩa tương đối, phụ thuộc vào đặc điểm của hiện tượng và từng loại chỉ tiêu khác nhau.
- Ví dụ: Sản phẩm của ngành chế biến thủy sản có thể xét theo ngày, tuần. Nhưng sản phẩm của ngành đóng tàu phải xét theo tháng, năm...



Hạn chế của phương pháp:

- Do ghép nhiều khoảng thời gian vào thành một nên số lượng các mức độ trong dãy số mất đi quá nhiều, đôi khi làm mất ảnh hưởng của các nhân tố cơ bản.
- Ví dụ: số liệu từ tháng chuyển thành quý, từ 12 mức độ còn 4 mức độ, tức là mất đi $\frac{2}{3}$ số mức độ ban đầu.
- Trường hợp sử dụng với những hiện tượng có tính chất thời vụ sẽ làm mất đi tính chất thời vụ của hiện tượng

2 – Phương pháp dãy số bình quân trượt (bq di động)

- Đặc điểm của số bình quân là san bằng các chênh lệch vì thế nó san bằng các nhân tố ngẫu nhiên làm bộc lộ nhân tố cơ bản của hiện tượng,
- Khái niệm: Số bình quân trượt là số bình quân của một nhóm nhất định các mức độ trong dãy số được tính bằng cách lần lượt loại trừ dần mức độ đầu, đồng thời thêm vào các mức độ tiếp theo sao cho số lượng các mức độ tham gia tính số bình quân là không đổi.
- Dãy số bình quân trượt là dãy số được hình thành từ các số bình quân trượt.

Công thức

- Ví dụ: Có dãy số thời gian n mức độ $y_1, y_2, \dots, y_n$ Giả sử nhóm 3 mức độ để tính số bình quân trượt, ta có:

$$\overline{y_2} = \frac{y_1 + y_2 + y_3}{3}$$

$$\overline{y_3} = \frac{y_2 + y_3 + y_4}{3}$$

$$\overline{y_{n-1}} = \frac{y_{n-2} + y_{n-1} + y_n}{3}$$

- $\overline{y_2}, \overline{y_3}, \dots, \overline{y_{n-1}}$ được gọi là dãy số bình quân trượt lần thứ nhất (MA_1).
- Nếu dãy số vẫn chưa bộc lộ rõ xu hướng, nghĩa là chưa loại bỏ hết các yếu tố ngẫu nhiên thì có thể tính bình quân trượt thêm một lần nữa.

- VD trên :
Tính số bình
quân trượt
theo nhóm 3
mức độ:

Tháng	Sản lượng (1000 tấn)(y_i)	Số bình quân trượt
1	40	-
2	36	39
3	41	38,33
4	38	40,33
5	42	42,67
6	48	43,33
7	40	44,33
8	45	45
9	50	48
10	49	48,33
11	46	45,67
12	42	

Chú ý:

- Tùy theo đặc điểm, tính chất của hiện tượng để xác định số các mức độ tham gia tính số bình quân trượt.
- Từ một dãy số có n mức độ, tính số bình quân trượt theo nhóm m mức độ thì số các mức độ của dãy số mới sẽ là $(n-m+1)$.

3 – Phương pháp hồi qui

- Trên cơ sở dãy số thời gian, XD phương trình hồi qui để biểu hiện xu hướng phát triển của hiện tượng. Dạng tổng quát của phương trình hồi qui theo thời gian (còn gọi là hàm xu thế):
- $y_t = f(t, a_0, a_1, \dots, a_n)$
- Với t là biến số thời gian. là thứ tự thời gian theo quy ước, đóng vai trò là biến số độc lập trong phương trình hồi quy.
- Các dạng hàm xu thế:
 - Tuyến tính
 - Parabol
 - Hypebol

Hàm xu thế tuyến tính:

- sử dụng khi dãy số thời gian có các lượng tăng (giảm) tuyệt đối liên hoàn xấp xỉ nhau.

Hàm có dạng: $\hat{y}_t = a_0 + a_1 t$

- Các tham số a_0 , a_1 được xác định bằng phương pháp bình phương nhỏ nhất. Theo đó, a_0 và a_1 phải thỏa mãn phương trình:

$$\begin{cases} \Sigma y = na_0 + a_1 \Sigma t \\ \Sigma ty = a_0 \Sigma t + a_1 \Sigma t^2 \end{cases}$$

hoặc: $a_1 = \frac{\overline{ty} - \bar{t}\bar{y}}{\sigma_t^2}$ và $a_0 = \bar{y} - a_1 \bar{t}$

Hàm xu thế parabol

Được sử dụng trong trường hợp các mức độ của dãy số tăng dần theo thời gian đạt cực đại, sau đó lại giảm dần theo thời gian hoặc giảm dần theo thời gian đạt cực tiểu, sau đó lại tăng dần theo thời gian

Hàm có dạng: $\hat{y}_t = a_0 + a_1t + a_2t^2$

$$\Sigma y = a_0.n + a_1. \Sigma t + a_2. \Sigma t^2$$

$$\Sigma ty = a_0. \Sigma t + a_1. \Sigma t^2 + a_2. \Sigma t^3$$

$$\Sigma t^2y = a_0. \Sigma t^2 + a_1. \Sigma t^3 + a_2. \Sigma t^4$$

Hàm xu thế Hypebol

- Hàm có dạng

$$\hat{y}_t = a_0 + \frac{a_1}{t}$$

- a_0 và a_1 được tìm theo hệ phương trình

$$\sum y = a_0 n + a_1 \sum \frac{1}{t}$$

$$\sum \frac{1}{t} y = a_0 \sum \frac{1}{t} + a_1 \sum \frac{1}{t^2}$$

Hàm xu thế mũ

- được sử dụng khi các tốc độ phát triển liên hoàn xấp xỉ nhau

Hàm có dạng: $\hat{y}_x = a_0 \cdot a_1^t$

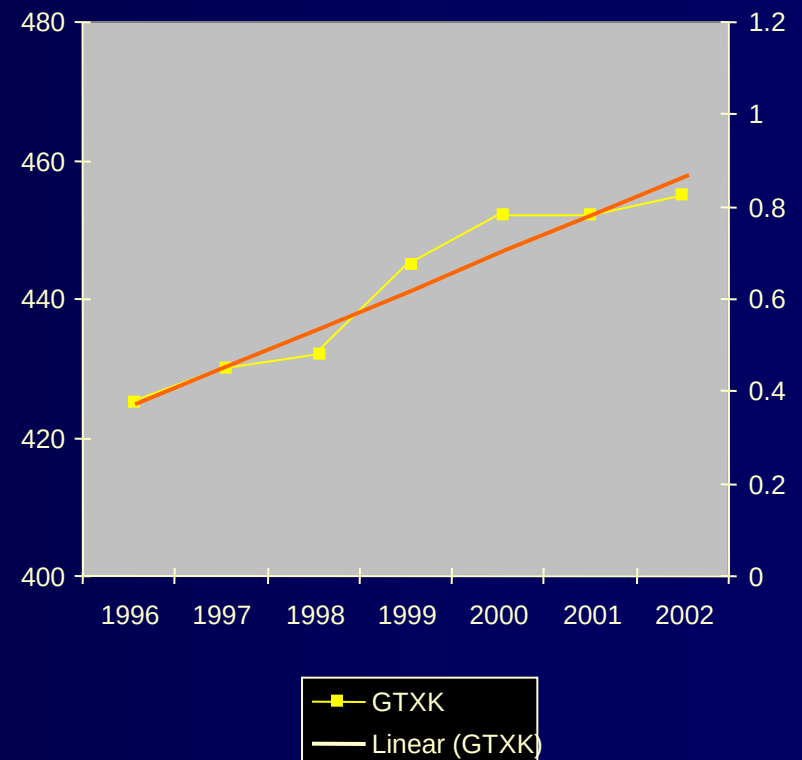
hay: $\ln y = \ln a_0 + t \cdot \ln a_1$

- Các hệ số tìm được từ giải hệ phương trình

$$\begin{cases} \Sigma \ln y &= \ln a_0 \cdot n + \ln a_1 \cdot \Sigma t \\ \Sigma t \ln y &= \ln a_0 \cdot \Sigma t + \ln a_1 \cdot \Sigma t^2 \end{cases}$$

Ví dụ:

Năm	GTXK (\$)
1996	425
1997	430
1998	432
1999	445
2000	452
2001	452
2002	455
Σ	???



Bảng số liệu

t	y	y.t	t^2
1	425	425	1
2	430	860	4
3	432	1296	9
4	445	1780	16
5	452	2260	25
6	452	2712	36
7	455	3185	49
Σ 28	3091	12518	140

Xác định giá trị tham số

$$\begin{cases} \sum y = na + b \sum t \\ \sum yt = a \sum t + b \sum t^2 \end{cases} \Rightarrow \begin{cases} 3091 = 7a + 28b \\ 12518 = 28a + 140b \end{cases}$$

$$\Rightarrow \begin{cases} a = 419,571 \\ b = 5,5 \end{cases}$$

■ Hàm xu thế:

$$y = 419,571 + 5,5t$$

Lựa chọn t như thế nào? $\Sigma t = 0$

$$\begin{cases} \sum y = na + b \sum t \\ \sum yt = a \sum t + b \sum t^2 \end{cases}$$

$$\Rightarrow \begin{cases} \sum y = na \\ \sum yt = b \sum t^2 \end{cases}$$

$$\Rightarrow \begin{cases} a = \frac{\sum y}{n} \\ b = \frac{\sum yt}{\sum t^2} \end{cases}$$

Xét cho ví dụ trên

Năm	GTXK	t'	yt'	t'^2
1996	425	-3	-1275	9
1997	430	-2	-860	4
1998	432	-1	-432	1
1999	445	0	0	0
2000	452	1	452	1
2001	452	2	904	4
2002	455	3	1365	9
Σ	3091	0	154	28

Xác định được các giá trị của a,b

$$\Rightarrow \begin{cases} a = \frac{\sum y}{n} \\ b = \frac{\sum yt}{\sum t^2} \end{cases}$$

$$\Rightarrow \begin{cases} a = \frac{3091}{7} = 441,51 \\ b = \frac{154}{28} = 5,5 \end{cases}$$

- Phương trình hàm xu thế có dạng:

$$y = 441,571 + 5,5t'$$

Nhận xét

Hàm xu thế theo t:

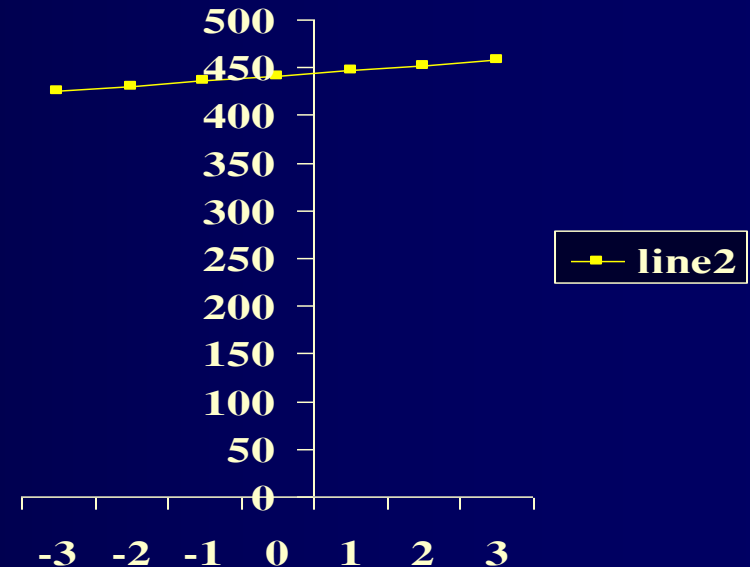
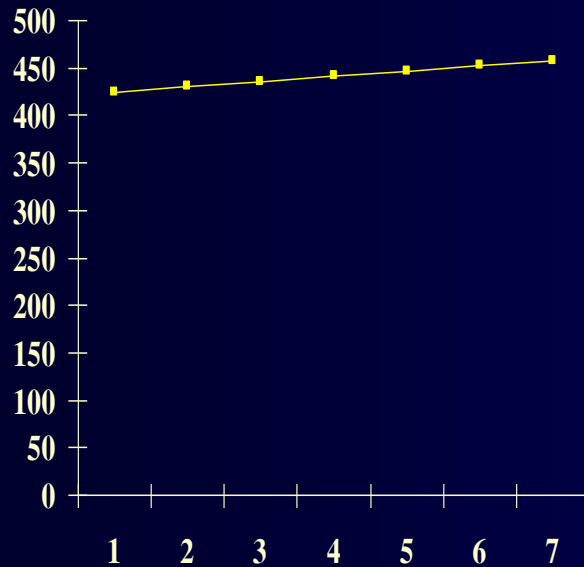
$$y = 419,571 + 5,5t$$

Hàm xu thế theo t'

$$y = 441,571 + 5,5t'$$

Kết quả ở hàm hồi quy sẽ khác nhau nhưng dùng để dự báo thì đều có giá trị như nhau.

Đồ thị hàm xu thế theo thời gian





**Nếu số lần thu thập số
liệu theo thời gian là số chẵn**

N''_m	GTXK (\$)	t	t'
1995	420	1	-7
1996	425	2	-5
1997	430	3	-3
1998	432	4	-1
1999	445	5	1
2000	452	6	3
2001	452	7	5
2002	455	8	7

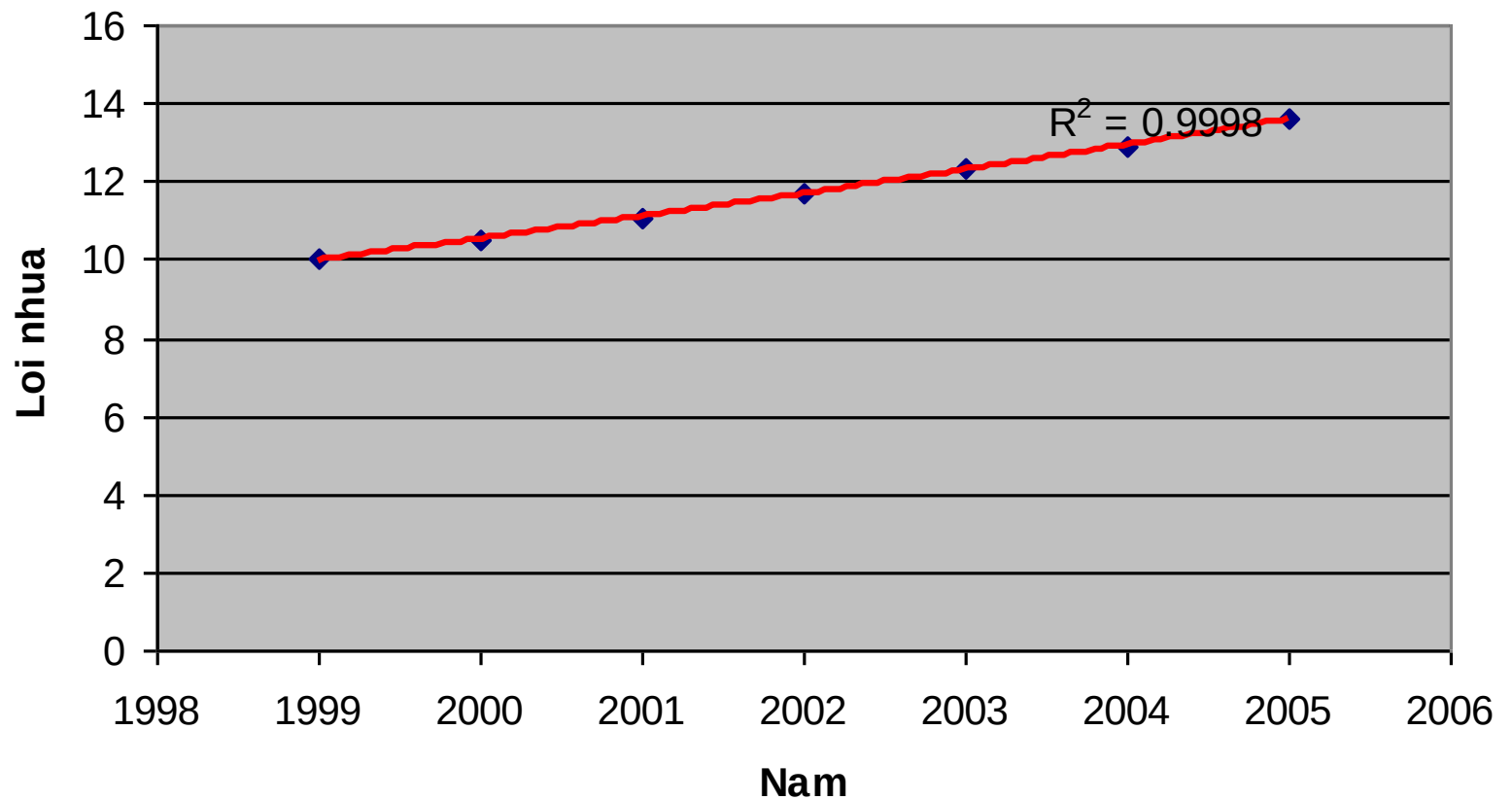
Bài tập

Năm	1999	2000	2001	2002	2003	2004	2005
Lợi nhuận (tỷ đồng)	10.0	10.5	11.1	11.7	12.3	12.9	13.6

Yêu cầu:

Xây dựng hàm xu thế theo thời gian

Đồ thị minh họa



Hàm xu thế tuyến tính có dạng:

$$\hat{y}_t = a_0 + a_1 t$$

Năm	Lợi nhuận (y)	Thứ tự thời gian (t)	ty	t ²
1999	10.00	1	10	1
2000	10.50	2	21	4
2001	11.10	3	33.3	9
2002	11.70	4	46.8	16
2003	12.30	5	61.5	25
2004	12.90	6	77.4	36
2005	13.60	7	95.2	49
Cộng	82.10	28	345.2	140
Trung bình	11.73	4.00	49.31	20.00

$$a_1 = \frac{\overline{ty} - \bar{t}\bar{y}}{\sigma_t^2} = \frac{49.31 - 4.0 \times 11.73}{20 - (4.0)^2} = 0.6$$

$$a_0 = \bar{y} - a_1 \bar{t}$$

$$= 11.73 - 0.6 \times 4.0 = 9.33$$

Hàm xu thế theo t: $y = 9,33 + 0.6t$

4. Nghiên cứu biến động thời vụ

■ a. Khái niệm

- Biến động lặp đi lặp lại của hiện tượng trong từng khoảng thời gian nhất định làm cho mức độ của nó lúc tăng, lúc giảm

■ Nguyên nhân:

- Do điều kiện tự nhiên
- Do tập quán sinh hoạt của dân cư.

- Biến động thời vụ thường ảnh hưởng đến hoạt động sản xuất kinh doanh của ngành đó và các ngành có liên quan. Việc nghiên cứu biến động thời vụ cho phép chủ động trong công tác quản lý kinh tế - xã hội, lập kế hoạch sản xuất hay hoạt động nghiệp vụ thích hợp, hạn chế ảnh hưởng đến sản xuất và sinh hoạt xã hội

b. Chỉ số thời vụ

- Để phản ánh biến động thời vụ, sử dụng chỉ số thời vụ
- Công thức:

$$I_i = \frac{\overline{y_i}}{\overline{y_0}} * 100\%$$

Trong đó: I_i : chỉ số thời vụ của thời gian thứ i (có thể là tháng, quý,...)

$\overline{y_i}$: mức độ bình quân của thời gian i qua các năm

$\overline{y_0}$: mức độ bình quân chung của dãy số

Ví dụ: Mức tiêu thụ hàng hóa trong 3 năm của doanh nghiệp A như sau:

Năm Quý	Mức tiêu thụ hàng hóa (trđ)			$\bar{y}_i$	I_i (%)
	2006	2007	2008		
I	4489	4589	4574	4551	63.86
II	7957	8296	8000	8084	113.46
III	9450	9524	9514	9496	133.27
IV	6376	6294	6444	6371	89.41
Σ	28272	28703	28532		400.00

Tính các mức độ bình quân của từng quý qua 3 năm $\bar{y}_i$

$$\bar{y}_1 = \frac{4489 + 4589 + 4574}{3} = 4551$$

Tính mức độ bình quân chung:

$$\bar{y}_0 = \frac{28272 + 28703 + 28532}{12} = 7126$$

Tính các chỉ số thời vụ cho từng quý I_i

$$I_1 = \frac{\bar{y}_1}{\bar{y}_0} \times 100 = \frac{4551}{7126} \times 100 = 63,8\%$$

Nhận xét: Mặt hàng này tiêu thụ mạnh (trên mức bình quân chung) trong quý II và quý III và tiêu thụ ít (dưới mức bình quân chung) vào các quý I, quý IV

Đối với dãy số có xu thế

- Nếu các mức độ cùng kỳ của hiện tượng từ năm này qua năm khác có biểu hiện tăng hay giảm rõ rệt, muốn tính chỉ số thời vụ, trước hết phải điều chỉnh dãy số bằng phương trình hồi quy để tính các mức độ lý thuyết rồi sau đó dùng các mức độ này làm căn cứ so sánh

- Công thức

$$I_i = \frac{\sum_{j=1}^n \frac{y_{ij}}{\hat{y}_{ij}}}{m} \times 100(\%)$$

- Trong đó: y_{ij} : mức độ thực tế của thời kỳ thứ i ($i=1, \dots, n$) thuộc năm j ($j=1, \dots, m$)
- $\hat{y}_{ij}$: mức độ lý thuyết của thời kỳ thứ i ($i=1, \dots, n$) thuộc năm j ($j=1, \dots, m$) được tính từ hàm xu thế
- m : số năm nghiên cứu

$$\frac{y_{ij}}{\hat{y}_{ij}} \times 100(\%)$$

Ví dụ

Quý	Mức độ thực tế (trđ) y_{ij}			Mức độ lý thuyết (trđ) $\hat{y}_{ij}$			$\frac{y_{ij}}{\hat{y}_{ij}} \times 100(\%)$			I_i (%)
	2006	2007	2008	2006	2007	2008	2006	2007	2008	
I	1639	2336	3030	866	1808	2750	189.28	129.21	110.18	142.89
II	864	1091	2177	1101	2043	2986	78.44	53.39	72.92	68.25
III	671	1407	2603	1337	2279	3221	50.19	61.74	80.81	64.25
IV	2410	2749	4958	1572	2515	3457	153.26	109.32	143.44	135.34
Σ	5584	7583	12768	4877	8645	12413				

IV. Một số phương pháp dự báo thống kê ngắn hạn

1. Khái niệm
2. Một số phương pháp thông dụng
 - Phương pháp sử dụng lượng tăng/giảm tuyệt đối bình quân
 - Phương pháp sử dụng tốc độ phát triển bình quân
 - Phương pháp ngoại suy hàm xu thế

Khái niệm

- Dự đoán thống kê là việc xác định các mức độ của hiện tượng nghiên cứu trong tương lai. Dự đoán chia ra làm 3 loại: dài hạn (>10 năm), trung hạn (3-10 năm) và ngắn hạn (<3 năm)
- Thống kê thường thực hiện dự đoán ngắn hạn
- Nội dung cơ bản của dự đoán thống kê là dựa trên các giá trị đã biết, phân tích các yếu tố ảnh hưởng đến sự biến động của hiện tượng, thừa nhận rằng những yếu tố sẽ vẫn còn tác động đến hiện tượng trong tương lai
- Có 3 phương pháp dự đoán cơ bản sau:
 - Phương pháp chuyên gia: tham khảo, hỏi ý kiến các chuyên gia
 - Phương pháp hồi quy: mô hình hoá hiện tượng chịu tác động của nhiều nhân tố từ đó xây dựng phương trình hồi quy.
 - Phương pháp dãy số thời gian: xây dựng mô hình và dự đoán

1. Phương pháp dự báo sử dụng lượng tăng/giảm tuyệt đối bình quân

- Phương pháp này được áp dụng khi lượng tăng hoặc giảm tuyệt đối liên hoàn của hiện tượng qua thời gian xấp xỉ bằng nhau.

$$y_{n+h} = y_n + \bar{\delta}h$$

h : Tầm xa của dự đoán

y_n : Mức độ cuối cùng của dãy số thời gian

$\bar{\delta}$: Lượng tăng giảm tuyệt đối bình quân

Dự báo giá trị xuất khẩu 2003 và 2004

Năm	1998	1999	2000	2001	2002
GTXK (\$)	40	43	45	49	52
Lượng T/G LH(\$)	-	3	2	4	3
Lượng T/G bq (\$)	3				

- Lượng tăng/giảm tuyệt đối bình quân trong cả giai đoạn là 3 (\$)/năm

$$\Rightarrow y_{2003} = 52 + 3 \cdot 1 = 55 (\$)$$

$$\Rightarrow y_{2004} = 52 + 3 \cdot 2 = 58 (\$)$$

2. Phương pháp dự báo sử dụng tốc độ phát triển bình quân

- Phương pháp này được áp dụng khi tốc độ phát triển liên hoàn của hiện tượng qua thời gian xấp xỉ bằng nhau

Công thức:

$$\hat{y}_{n+h} = y_n \left(\frac{-}{t} \right)^h$$

- h: Tầm xa của dự đoán
- y_n : Mức độ cuối cùng trong dãy số thời gian
- t: Tốc độ phát triển bình quân

Dự báo GTXK cho năm 2003 và 2004

Năm	1998	1999	2000	2001	2002
GTXK (\$)	40	43	45	49	52
Tốc độ pt LH (lần)	-	1,075	1,047	1,089	1,061
Tốc độ pt bq (lần)	1,0678				

- Tốc độ phát triển bình quân trong cả giai đoạn là 106,78 %/năm
- $y_{2003} = 52 * 1,0678^1 = 55,53$ (\$)
- $y_{2004} = 52 * 1,0678^2 = 59,29$ (\$)

3. Phương pháp ngoại suy hàm xu thế

- Phương pháp này dựa trên hàm hồi quy biểu diễn xu thế phát triển của hiện tượng.

- Ta có hàm xu thế :

$$y_t = f(t, a_0, a_1, a_2, \dots, a_n)$$

- Giá trị dự đoán: thay giá trị $t+h$ tương ứng với thời kỳ cần dự đoán để tìm ra:

$$y_{t+h} = f(t+h, a_0, a_1, a_2, \dots, a_n)$$

Ví dụ: Có tài liệu về một DN như sau:

Năm	98	99	00	01	02	03	04	05
TSCĐ (tỷ VND)	80	87	95	102	111	121	130	140

- Lập phương trình hồi quy biểu diễn xu thế phát triển của quy mô TSCĐ
- Hãy dự đoán quy mô TSCĐ năm 2006

4. Dự đoán kết hợp xu thế và biến động thời vụ

- Mô hình dự đoán

$$\hat{y}_t = f(t) \times I_i$$

- với $f(t)$ là hàm xu thế tìm được, I_i là chỉ số thời vụ đối với từng thời kỳ cụ thể, thay giá trị t tương ứng để tìm $\hat{y}_t$.

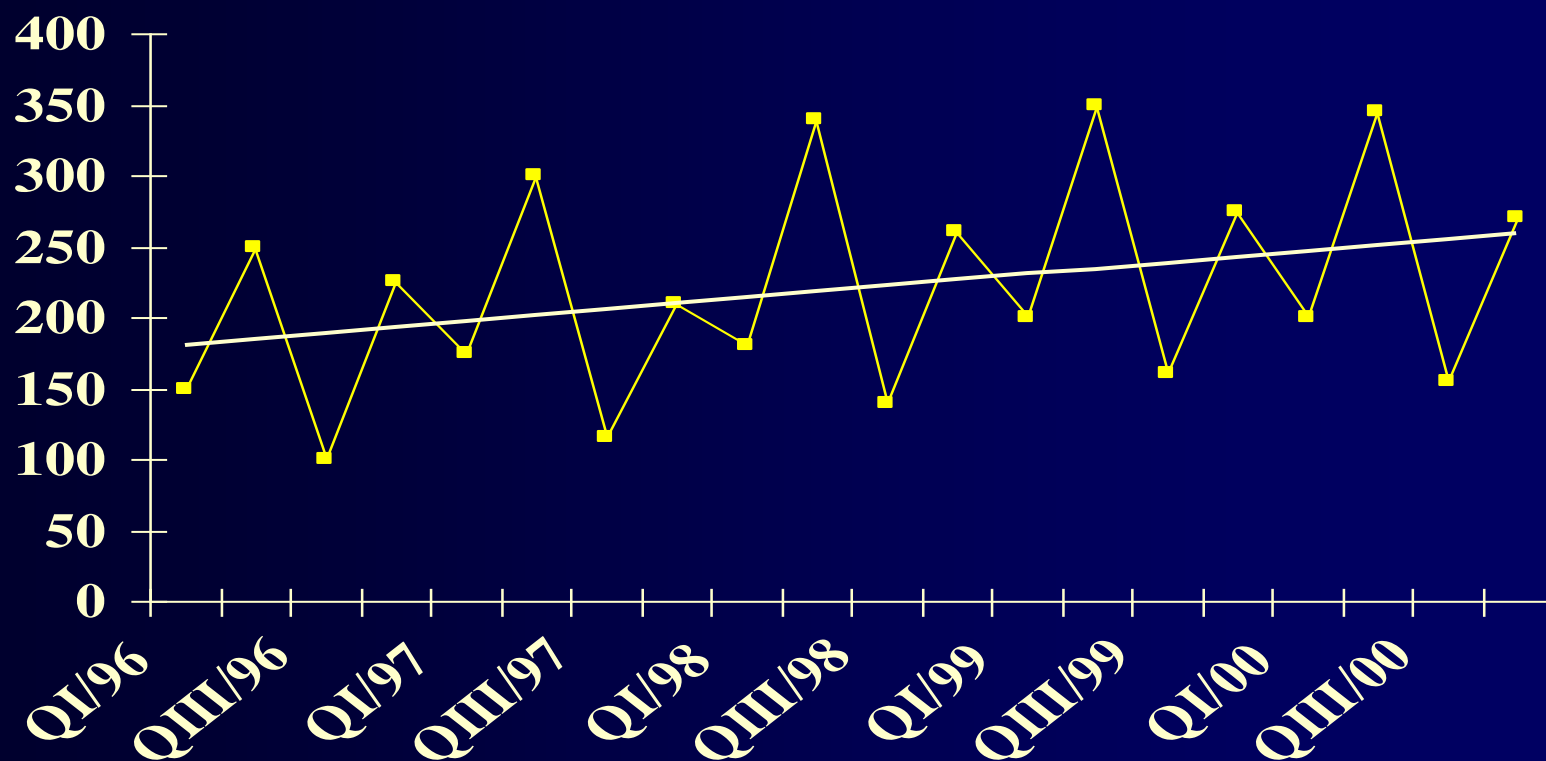
Bài tập: số liệu theo dõi lượng du khách đến Sapa

(đv: nghìn lượt). Yêu cầu:

- Phân tích biến động thời vụ
- Dự báo lượng khách đến theo mùa năm 2006

	1996	1997	1998	1999	2000	$\bar{y}_i$	I_i
Mùa xuân	150	175	180	200	200	181	82.27
Mùa hạ	250	300	340	350	345	317	144.09
Mùa thu	100	115	140	160	155	134	60.91
Mùa đông	225	210	260	275	270	248	112.73
bq	220						100

Đường biểu diễn xu thế biến động thời vụ



Kết hợp nghiên cứu biến động thời vụ và dự báo bằng phương pháp ngoại suy để xác định mức độ thời vụ trong tương lai

Dự báo cho năm 2006

Năm	t'	y	t'^2	yt'
2000	-2	725	4	-1450
2001	-1	800	1	-800
2002	0	920	0	0
2003	1	980	1	980
2004	2	975	4	1950
Σ	0	4400	10	680

Kết hợp nghiên cứu biến động thời vụ và dự báo bằng phương pháp ngoại suy để xác định mức độ thời vụ trong tương lai

$$a = \frac{\sum y}{n} = \frac{4400}{5} = 880$$

$$b = \frac{\sum yt'}{\sum t'^2} = \frac{680}{10} = 68$$

- Hàm xu thế: $y = 880 + 68t'$
- Năm 2006 $\rightarrow t' = 4$
- $y_{06} = 880 + 68 \cdot 4 = 1152$ (ngL)
- $y_{QI/06} = (1152/4) \cdot 82,27\% = 236,94$ (nghìn lượt)
- $y_{QII/06} = \dots$

Bài tập: Mức lưu chuyển mặt hàng fast-food tại 1 cửa hàng

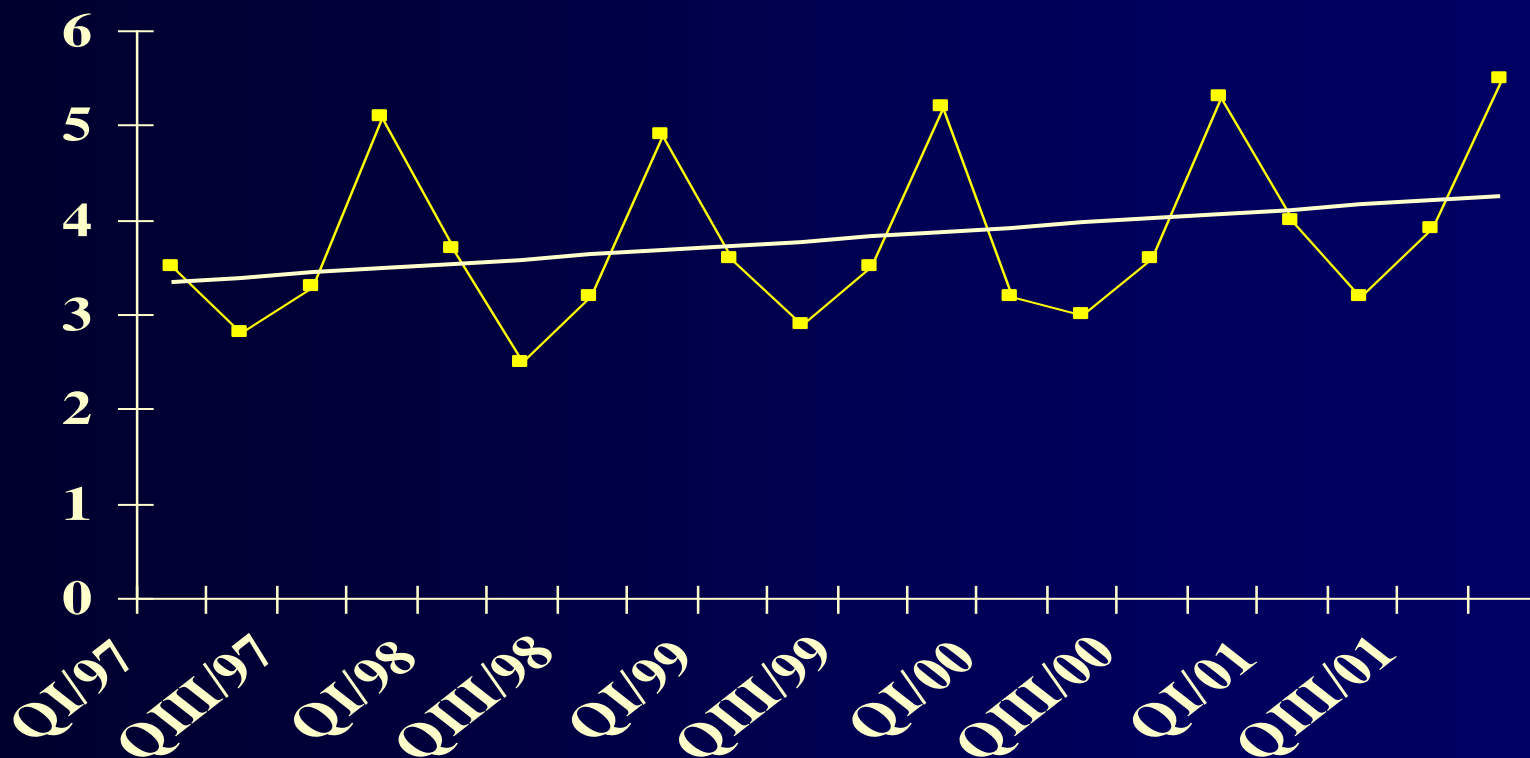
	2001	2002	2003	2004	2005
Xuân	3.5	3.7	3.6	3.2	4.0
Hạ	2.8	2.5	2.9	3.0	3.2
Thu	3.3	3.2	3.5	3.6	3.9
Đông	5.1	4.9	5.2	5.3	5.5

Phân tích biến động thời vụ về lưu chuyển tiền mặt
Dự báo ngắn hạn lưu chuyển tiền mặt từng quý
năm 2007

Phân tích biến động thời vụ

	2001	2002	2003	2004	2005	$\bar{x}_i$	$I_i (\%)$
Xuân	3.5	3.7	3.6	3.2	4.0	3.60	94.86
Hạ	2.8	2.5	2.9	3.0	3.2	2.88	75.89
Thu	3.3	3.2	3.5	3.6	3.9	3.50	92.23
Đông	5.1	4.9	5.2	5.3	5.5	5.20	137.02
						3.795	

Đường biểu diễn xu thế biến động thời vụ



Kết hợp nghiên cứu biến động thời vụ và dự báo bằng phương pháp ngoại suy để xác định mức độ thời vụ trong tương lai

Dự báo

Năm	t'	y	yt'	t'^2
2001	-2	14,7	-29,4	4
2002	-1	14,3	-14,3	1
2003	0	15,2	0	0
2004	1	15,1	15,2	1
2005	2	16,6	33,2	4
		75,9	4,7	10

Kết hợp nghiên cứu biến động thời vụ và dự báo bằng phương pháp ngoại suy để xác định mức độ thời vụ trong tương lai

$$a = \frac{\sum y}{n} = \frac{75,9}{5} = 15,18$$

$$b = \frac{\sum yt'}{\sum t'^2} = \frac{4,7}{10} = 0,47$$

- Hàm xu thế: $y = 15,18 + 0,47t'$
- Năm 2007 $\rightarrow t' = 4$
- $Y_{07} = 15,18 + 0,47 \cdot 4 =$
- Dự báo cho mùa xuân

Bài tập 6/355