

TRƯỜNG ĐẠI HỌC BÁCH KHOA TP. HCM
 Khoa KTXD - Bộ môn KTTNN

PHÂN TÍCH DỮ LIỆU



Giảng viên: PGS. TS. NGUYỄN THÔNG
 E-mail: nguyenthong@hcmut.edu.vn or nthong56@yahoo.fr
 Web: <http://www4.hcmut.edu.vn/~nguyenthong/>
 PGS. TS. Nguyễn Thông | Tel: (08) 38 691 592- 098 99 66 719

PHÂN TÍCH DỮ LIỆU
 Chương 7: Phân tích chuỗi thời gian

Chương 1. Thống kê mô tả (ôn).
Chương 1bis. Xác suất & phân phối thống kê (ôn).
Chương 2. Khoảng tin cậy.
Chương 3. Kiểm định thống kê.
Chương 4. Phân loại dữ liệu (Classification).
Chương 5. Phân nhóm dữ liệu (Cluster).
Chương 6. Phân tích thành phần chính (PCA).
Chương 7. Phân tích chuỗi thời gian.
Chương 8. Hồi quy tuyến tính.
Chương 9. Xử lý số liệu thực nghiệm.
Chương 10. Giới thiệu phần mềm SPSS or R

PGS. TS. Nguyễn Thông

PHÂN TÍCH DỮ LIỆU
 Chương 7: Phân tích chuỗi thời gian

PHÂN TÍCH CHUỖI THỜI GIAN

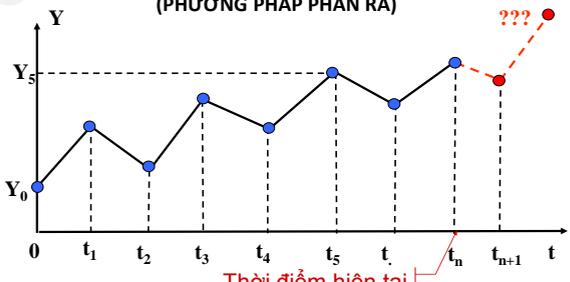
→ Phương pháp phân rã
 (Decomposition method)

→ Phương pháp Box-Renkins
 (AR → Auto Regressive
 MA → Mobile Average
 ARMA, ARIMA, SARIMA)

PGS. TS. Nguyễn Thông

PHÂN TÍCH DỮ LIỆU
 Chương 7: Phân tích chuỗi thời gian

PHÂN TÍCH CHUỖI THỜI GIAN (PHƯƠNG PHÁP PHÂN RÃ)



Thời điểm hiện tại

PGS. TS. Nguyễn Thông

PHÂN TÍCH DỮ LIỆU
 Chương 7: Phân tích chuỗi thời gian

PHƯƠNG PHÁP PHÂN RÃ

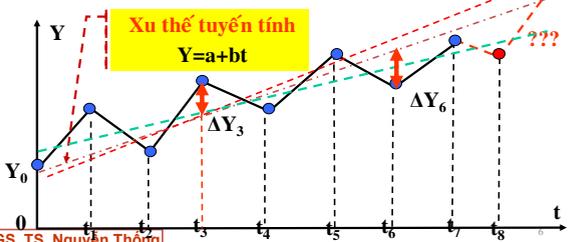
Chuỗi số liệu được nghiên cứu tách biệt theo 2 yếu tố:

→ Xu thế số liệu, trend (vĩ mô)
 → Biến đổi mùa, season (vi mô)

PGS. TS. Nguyễn Thông

PHÂN TÍCH DỮ LIỆU
 Chương 7: Phân tích chuỗi thời gian

1. PHÂN TÍCH XU THẾ: Đây là một phân tích liên quan đến chuỗi nhiều năm, do đó ta sẽ sử dụng số liệu hàng năm để phân tích. Một cách tổng quát ta cần phải có một chuỗi dài ít ra là 10 → 15 năm.



Xu thế tuyến tính
 $Y=a+bt$

PGS. TS. Nguyễn Thông

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

- Để đánh giá yếu tố xu thế, phương pháp sử dụng phổ biến là **phương pháp bình phương tối thiểu (BPTT)**.
- Đây là phương pháp cho phép xác định được đường cong (thẳng) hoặc mặt phẳng (siêu mặt phẳng) biểu thị xu thế số liệu, giới thiệu "tốt nhất" số liệu trong quá khứ ("gần với số liệu quan sát").

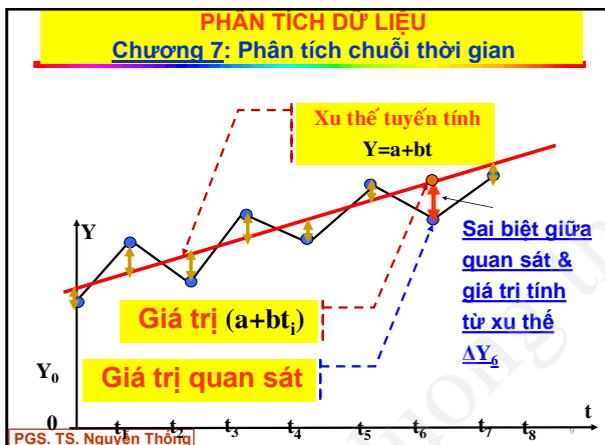
PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

- Trong trường hợp cá biệt khi nhận thấy xu thế của biến khảo sát trong thời gian dài là **tuyến tính**, phương trình sẽ xác định bởi:

$$Y = a + bt$$

- trong đó t biểu thị thời gian (năm) và a, b chỉ các thông số được xác định đường thẳng tính được từ phương pháp BPTT.

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian



PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

- Gọi ΔY_i là khoảng cách thẳng đứng từ điểm quan sát (t_i, Y_i) đến đường thẳng cần xác định. Ta định nghĩa hàm mục tiêu:

$$D = \sum_{i=1}^N \Delta Y_i^2 = \sum_{i=1}^N [Y_i - (a + bt_i)]^2 \Rightarrow \min$$

- Đây là một hàm 2 biến a và b, để cho D cực trị (với ý nghĩa vật lý của bài toán ta biết đó là cực tiểu) ta phải có:

$$\begin{cases} \frac{\partial D}{\partial a} = 0 \\ \frac{\partial D}{\partial b} = 0 \end{cases}$$

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

Từ đó: $\sum_i 2[y_i - (a + bt_i)] = 0$ [1]

$\sum_i 2[y_i - (a + bt_i)] \cdot t_i = 0$ [2]

Giải hệ phương trình trên ta có:

$$b = \frac{\sum_i t_i y_i - N \cdot \bar{t} \cdot \bar{y}}{\left(\sum_i t_i^2\right) - N \cdot \bar{t}^2}$$

$$a = \bar{y}_i - b \bar{t}_i$$

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

$$\bar{t} = \frac{\sum_{i=1}^N t_i}{N} = \frac{t_1 + \dots + t_N}{N}$$

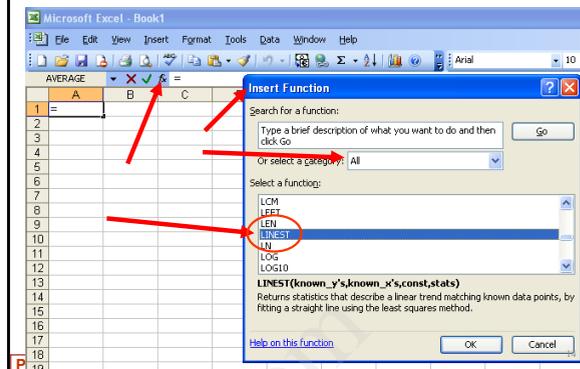
N → số quan trắc

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

→ **Intercept(Y,X)** → tính tung độ gốc a của quan hệ tuyến tính.

→ **Linest(Y,X)** → tính hệ số b (độ dốc của quan hệ tuyến tính giữa Y_i & X_j).

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian



PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

Chú ý: Trong trường hợp xu thế không phải là tuyến tính, ta có thể xét đến dạng đường cong hàm mũ $y = ab^t$ hoặc dạng parabol $y = a + bt + ct^2$.

→ Các thông số a, b, c vẫn xác định dựa vào khái niệm bình phương tối thiểu mà ta vừa nghiên cứu trên.

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian



ĐÁNH GIÁ
SỰ BIẾN ĐỔI
THEO MÙA

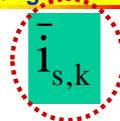
PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

Để nhận biết ảnh hưởng của thành phần mùa lên chuỗi thời gian khảo sát ta dùng thông số gọi là **chỉ số mùa**.

- Nếu số liệu theo **tháng**, ta có **12** giá trị i_s .
- Nếu số liệu theo **quý**, ta có **4** giá trị i_s .
- Nếu giá trị tính theo **6 tháng**, ta có **2** giá trị i_s .

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

TÍNH CHỈ SỐ MÙA



Từ số liệu quan trắc chỉ số mùa được tính từ $i_{s,t}$ như sau:

$$i_{s,t} = [\text{giá trị quan trắc}] / [\text{giá trị cho bởi } y = [a + bt]_t]$$

→ **Chú ý:** Có bao nhiêu số liệu quan trắc → có bấy nhiêu chỉ số $i_{s,t}$

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

Chú ý: Giá trị $i_{s,t}$ thay đổi chung quanh giá trị 1

$i_{s,t3} = AB/AC > 1$
 $i_{s,t4} < 1$

$Y \rightarrow i_{s,t}$ thay đổi chung giá trị 1

Xu thế tuyến tính
 $Y = a + bt$

PGS. TS. Nguyễn Thống

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

Từ các giá trị $i_{s,t}$, các giá trị đại biểu i_s được tính bằng giá trị trung bình của các tháng (quý) tương ứng:

- Theo tháng:
$$i_{s,k} = \frac{\sum_{t=1}^N i_{s,t=k}}{N}; k = 1,12$$
- Theo quý:
$$i_{s,k} = \frac{\sum_{t=1}^N i_{s,t=k}}{N}; k = 1,4$$

Với N (số tháng, quý,...) có trong chuỗi số liệu phân tích.

PGS. TS. Nguyễn Thống

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

GIÁ TRỊ CHỈ SỐ MÙA HIỆU CHỈNH

Ta phải có:

$\sum_{i=1}^{12} i_{s,i} = 12$ (tháng) $\sum_{i=1}^4 i_{s,i} = 4$ (quý)

\rightarrow Hiệu chỉnh:

$\bar{i}_{s,k} = \frac{12}{\sum_{i=1}^{12} i_{s,i}} i_{s,k}$ (tháng) $\bar{i}_{s,k} = \frac{4}{\sum_{i=1}^4 i_{s,i}} i_{s,k}$ (quý)

PGS. TS. Nguyễn Thống

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

Chuỗi CVS (loại bỏ ảnh hưởng mùa trong chuỗi giá trị quan sát)

Sự hiệu chỉnh mùa này cho phép chúng ta muốn so sánh kết quả của các tháng khác nhau trong một mùa nhằm để biết nếu có sự tăng hay giảm đã xảy ra so với giá trị bình thường. Giá trị hiệu chỉnh mùa sẽ được tính như sau:

$$y_{t,k}^* = \frac{y_t}{i_{s,k}} \rightarrow y_{t,k}^* \rightarrow \text{chuỗi CVS}$$

$k=1,12$ (số liệu tháng) hay $1,4$ (số liệu quý)

PGS. TS. Nguyễn Thống

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

DỰ BÁO VỚI MÔ HÌNH PHÂN RÃ

Giá trị dự báo tại thời điểm t của biến nghiên cứu y_t^* được xác định như sau:

$$y_t^* = [a + bt] * \bar{i}_{s,k}$$

Giá trị cho bởi đường xu thế

k tương ứng với tháng (mùa) tại thời điểm t.

PGS. TS. Nguyễn Thống

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

DỰ BÁO **ĐƯỜNG XU THẾ $y = a + bt$**

Giá trị dự báo tại $t=t_8$

Hiện tại

PGS. TS. Nguyễn Thống

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

Ví dụ: Số liệu kinh doanh theo quý (CA, tỷ đồng) của một Công ty cho 6 năm gần nhất như sau:

Quý	1	2	3	4
Năm 1	21	30	35	55
2	24	36	40	75
3	31	40	50	80
4	33	40	55	83
5	36	43	60	85
6	46	50	?	?

PGS. TS. Nguyễn Thống

25

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

SỐ LIỆU DƯỚI DẠNG ĐỒ THỊ



PGS. TS. Nguyễn Thống

26

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

a. Xác định xu thế biểu diễn bởi đường thẳng bằng phương pháp bình phương tối thiểu (a,b):

$$Y = a + bt$$

- b. Xác định hệ số biến đổi mùa đại biểu cho 4 quý.
c. Xác định doanh thu tính từ mô hình (quý 1/năm 1 đến quý 2/năm 6).
d. Hãy dự báo doanh thu cho quý 3 và 4 của năm cuối cùng và quý 1, 2 năm kế tiếp.

PGS. TS. Nguyễn Thống

27

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

HƯỚNG DẪN

PGS. TS. Nguyễn Thống

28

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

Quý	t	CA	Xu thế	i_{st}	Chỉ số mùa	Dự báo
(1)	(2)	(3)	(4)=bt+a	(5)=(3)/(4)	(6)	(7)=(4)*(6)
1	0	21	33.49	0.627	0.656	21.97
2	1	30	34.84	0.861	0.813	28.33
3	2	35	36.18	0.967	0.994	35.99
4	3	55	37.53	1.465	1.536	57.66
1	4	24	38.88	0.617	0.656	25.51
2	5	36	40.23	0.895	0.813	32.72
3	6	40	41.57	0.962	0.994	41.34
4	7	75	42.92	1.747	1.536	65.93

PGS. TS. Nguyễn Thống

29

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

1	8	31	44.27	0.7	0.656	29.04
2	9	40	45.62	0.877	0.813	37.1
3	10	50	46.96	1.065	0.994	46.7
4	11	80	48.31	1.656	1.536	74.21
1	12	33	49.66	0.665	0.656	32.58
2	13	40	51	0.784	0.813	41.48
3	14	55	52.35	1.051	0.994	52.06
4	15	83	53.7	1.546	1.536	82.49

PGS. TS. Nguyễn Thống

30

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

1	16	36	55.05	0.654	0.656	36.11
2	17	43	56.39	0.762	0.813	45.86
3	18	60	57.74	1.039	0.994	57.42
4	19	85	59.09	1.439	1.536	90.77
1	20	46	60.44	0.761	0.656	39.65
2	21	50	61.78	0.809	0.813	50.25
3	22	?	63.13		0.994	62.78
4	23	?	64.48		1.536	99.05

a=	33.49	Quý 1	0.671	0.656
b=	1.347	Quý 2	0.832	0.813
		Quý 3	1.017	0.994
		Quý 4	1.571	1.536
		Total	Sum=4,09	Sum=4,00

PGS. TS. Nguyễn Thống

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

DỰ BÁO

Quý 3 năm 6: $t=22$:
 $\rightarrow W_{22}=[33.49+1.347*22]*0.994 = 62.78$

Quý 4 năm 6: $t=23$:
 $\rightarrow W_{23}=[33.49+1.347*23]*1.536 = 99.05$

Quý 1 năm 7: $t=24$:
 $\rightarrow W_{24}=[33.49+1.347*24]*0.656 = 43.18$

Quý 2 năm 7: $t=25$:
 $\rightarrow W_{25}=[33.49+1.347*25]*0.813 = 54.60$

PGS. TS. Nguyễn Thống

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

Bài tập: Số liệu kinh doanh theo quý (CA, tỷ đồng) của một Công ty cho 5 năm gần nhất như sau:

Năm					
Quý	1	2	3	4	5
I	12	20	25	30	50
II	15	40	45	50	60
III	40	50	60	70	80
IV	60	75	90	95	110

PGS. TS. Nguyễn Thống



PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

a. Xác định xu thế (a,b) & xác định hệ số biến đổi mùa đại biểu cho 4 quý.
 b. Xác dự báo doanh thu tính từ mô hình (quý 1/năm 6 đến quý 4/năm 6).

HƯỚNG DẪN: Xem BAITAP-Chuoi thoi gian.xls

a=	23.73	$i_{s,t}$	$i_{s,t}$ hiệu chỉnh
b=	3.171	0.544	0.547
		0.794	0.798
		1.120	1.125
		1.523	1.530
		Sum=3.982	Sum=4.00

PGS. TS. Nguyễn Thống

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

Bài tập: Một Công ty kinh doanh nước ghi nhận nước tiêu thụ ($W, 10^6 m^3$) theo quý cho 6 năm gần đây như sau:

Quý	1	2	3	4
Năm 1	25	30	35	65
2	28	36	45	85
3	31	40	50	90
4	35	48	65	103
5	38	50	72	115
6	46	57	80	120

PGS. TS. Nguyễn Thống

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

- a. Giả thiết xu thế dạng đường thẳng, xác định phương trình bằng phương pháp bình phương tối thiểu ($Y=a+bt$).
- Từ đó xác định hệ số biến đổi mùa đại biểu cho 4 quý.
- b. Dự báo nhu cầu tiêu thụ từ mô hình cho quý 1,2,3 và 4 của năm kế tiếp.

PGS. TS. Nguyễn Thống

37

PHÂN TÍCH DỮ LIỆU
Chương 7: Phân tích chuỗi thời gian

HẾT
Xin cảm ơn!

PGS. TS. Nguyễn Thống

38