

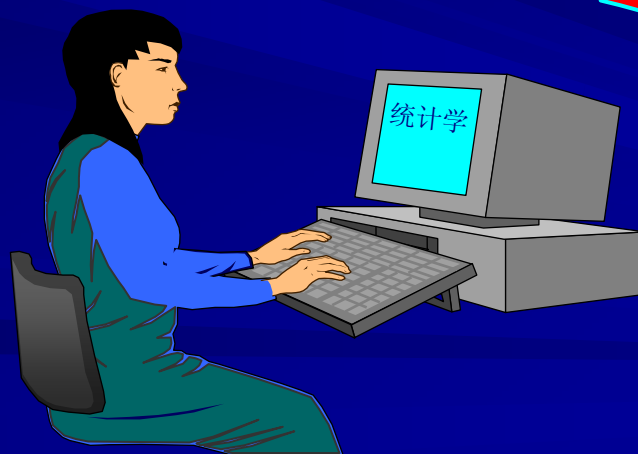
统计学

STATISTICS

(第三版)

第13章 时间序列分析和预测

PowerPoint



作者：张占贞，青岛科技大学经济与管理学院

第13章 时间序列分析和预测

- § 13.1 时间序列及其分解
- § 13.2 时间序列的描述性分析
- § 13.3 时间序列的预测程序
- § 13.4 平稳序列的预测
- § 13.5 趋势型序列的预测
- § 13.6 复合型序列的分解预测

学习目标

1. 时间序列及其分解原理
2. 时间序列的描述性分析
3. 时间序列的预测程序
4. 平稳序列的预测方法
5. 有趋势成分的序列的预测方法
6. 有季节成分的序列的预测方法
7. 复合型序列的分解预测

§ 13.1 时间序列及其分解

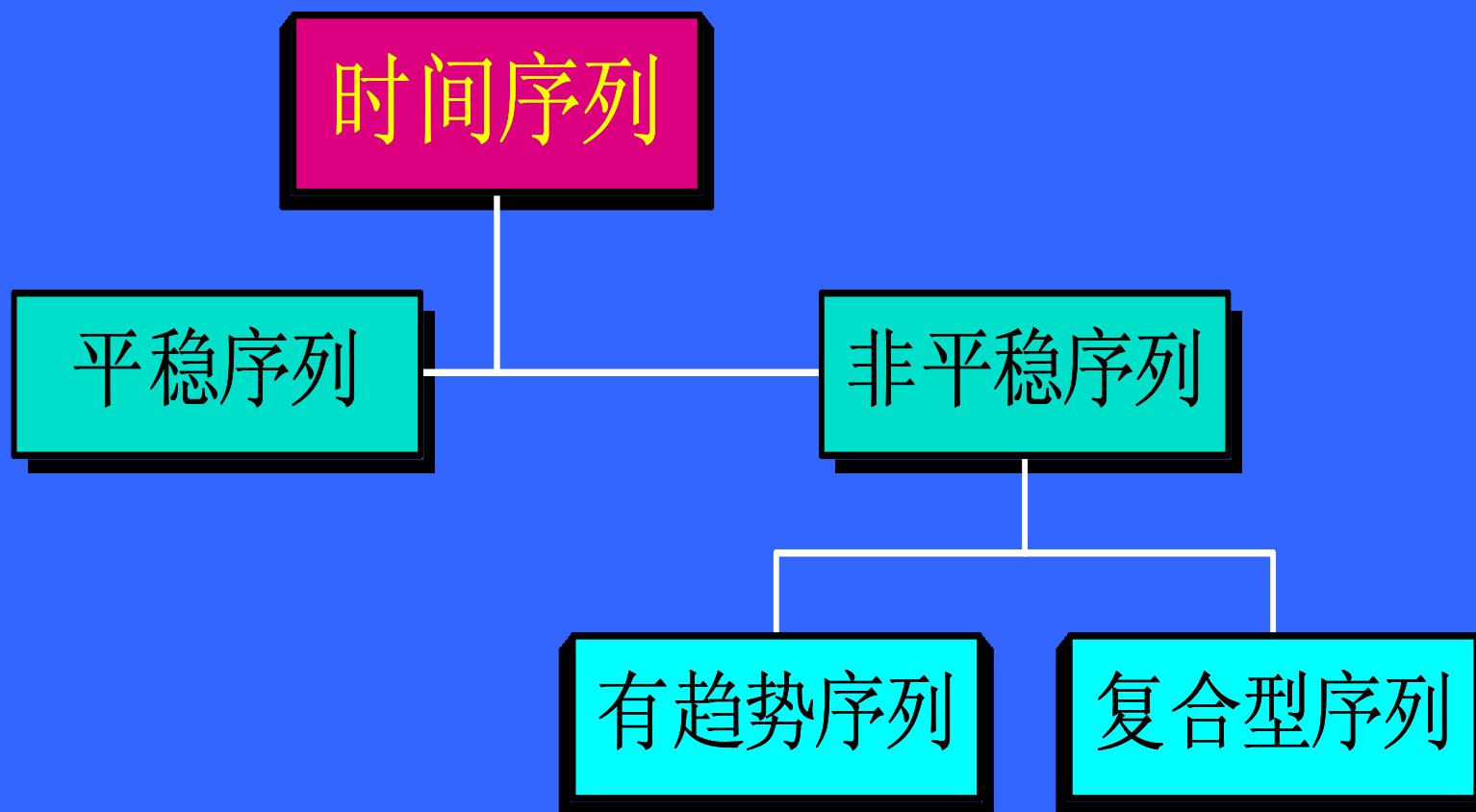
13.1.1 时间序列的构成要素

13.1.2 时间序列的分解方法

时间序列 (times series)

1. 同一现象在不同时间上的相继观察值排列而成的数列
2. 形式上由现象所属的时间和现象在不同时间上的观察值两部分组成
3. 排列的时间可以是年份、季度、月份或其他任何时间形式

时间序列的分类



时间序列的分类

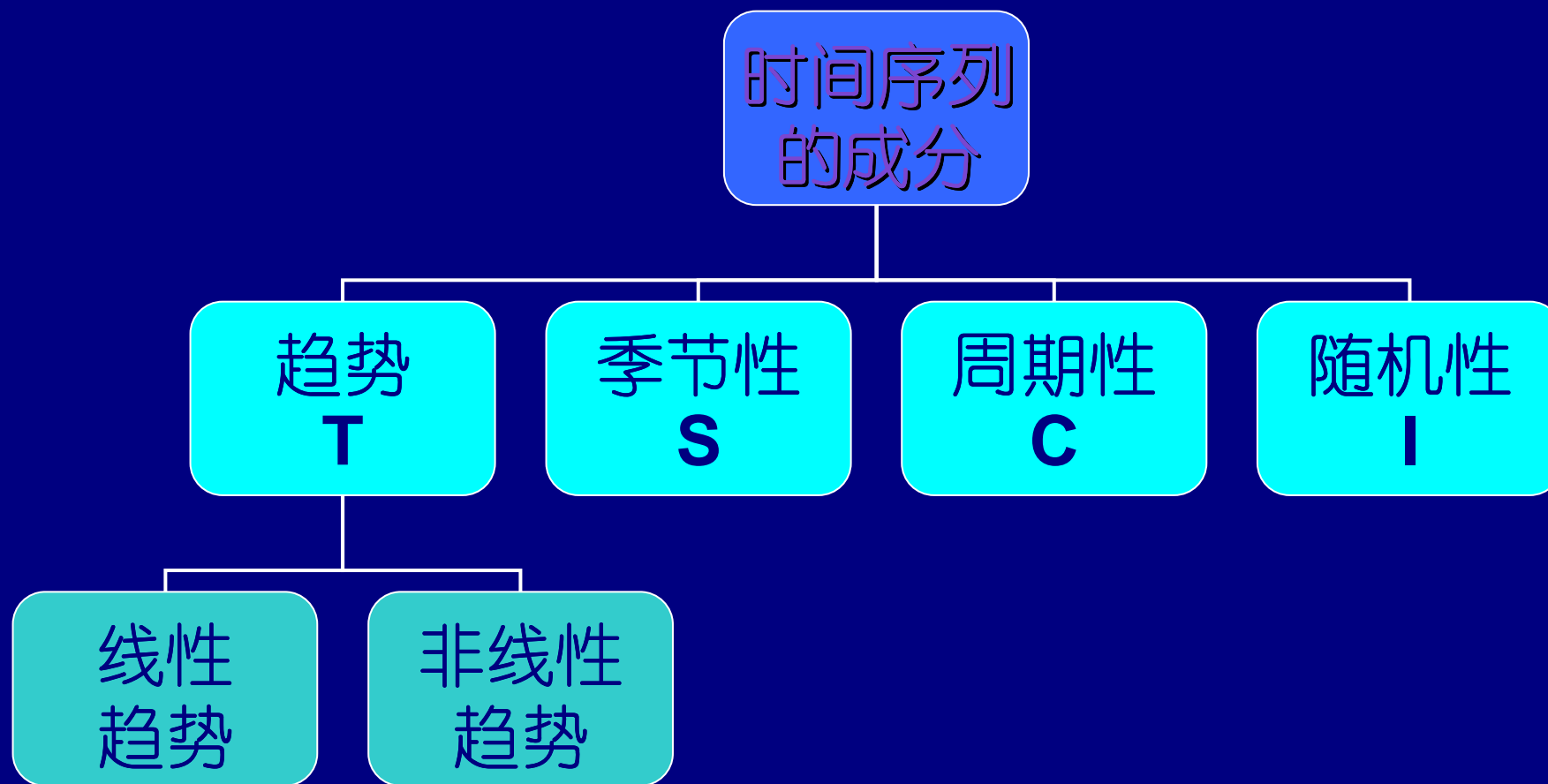
1. 平稳序列(stationary series)

- 基本上不存在趋势的序列，各观察值基本上在某个固定的水平上波动
- 或虽有波动，但并不存在某种规律，而其波动可以看成是随机的

2. 非平稳序列 (non-stationary series)

- 有趋势的序列
 - 线性的，非线性的
- 有趋势、季节性和周期性的复合型序列

时间序列的成分



时间序列的成分

1. 趋势(trend)

- 持续向上或持续下降的状态或规律

2. 季节性(seasonality)

- 也称季节变动(Seasonal fluctuation)
- 时间序列在一年内重复出现的周期性波动

3. 周期性(cyclity)

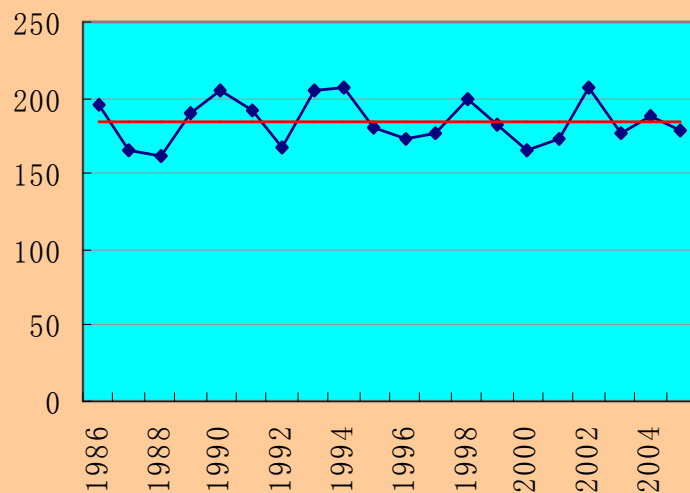
- 也称循环波动(Cyclical fluctuation)
- 围绕长期趋势的一种波浪形或振荡式变动

4. 随机性(random)

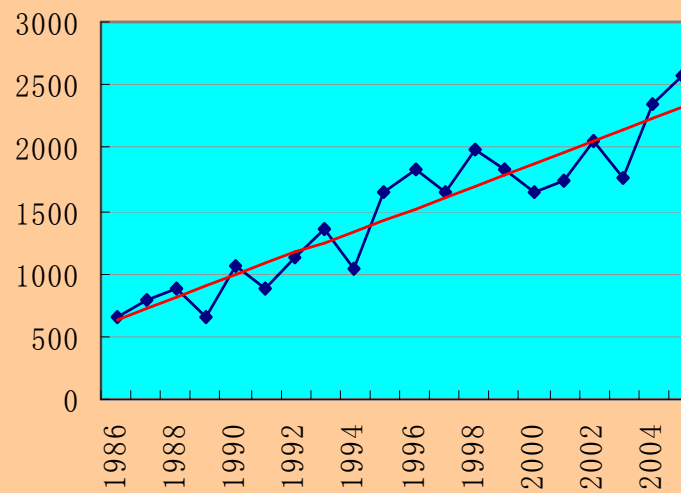
- 也称不规则波动(Irregular variations)
- 除去趋势、周期性和季节性之后的偶然性波动

含有不同成分的时间序列

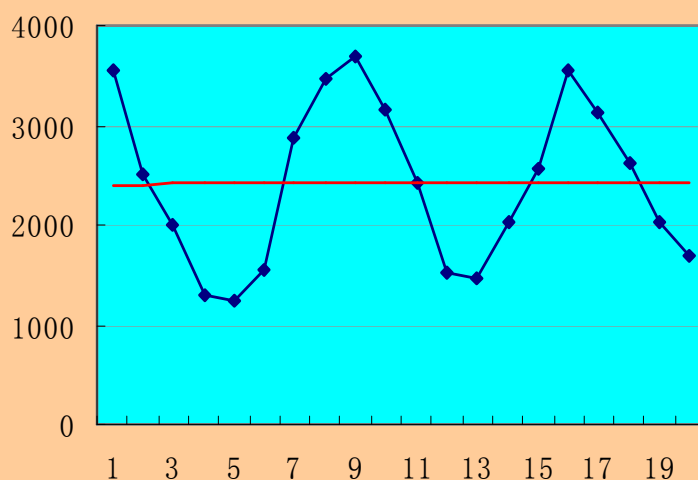
平稳



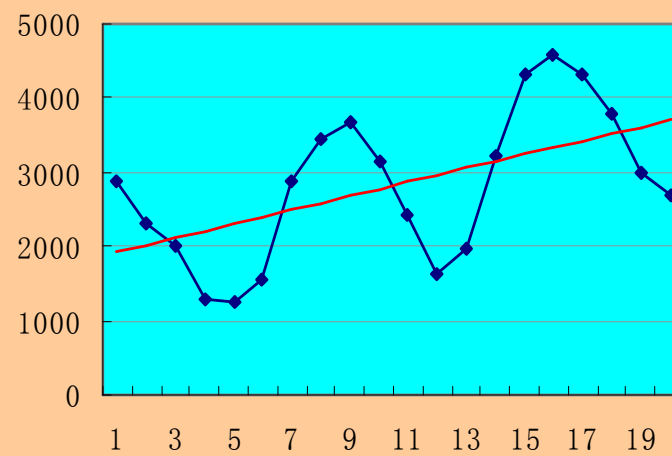
趋势



季节



季节与趋势



§13.2 时间序列的描述性分析

13.2.1 图形描述

13.2.2 增长率分析

统计学
STATISTICS
(第三版)

图形描述

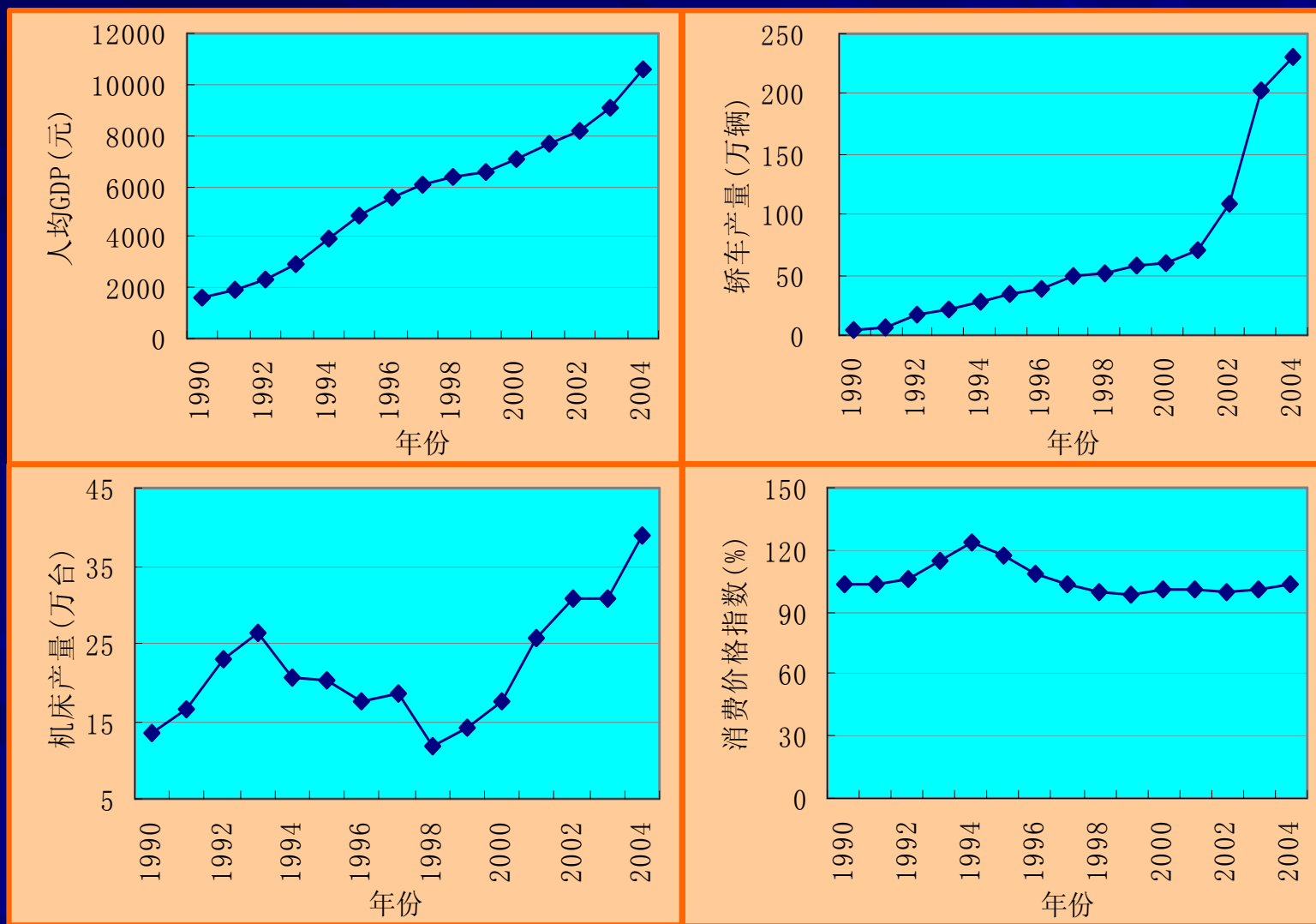
作者：张占贞，青岛科技大学经济与管理学院

图形描述 (例题分析)

	A	B	C	D	E
1	年份	人均GDP (元)	轿车产量 (万辆)	金属切削机床产量 (万台)	居民消费价格指数(%) (上年=100)
2	1990	1634	3.50	13.5	103.1
3	1991	1879	6.87	16.4	103.4
4	1992	2287	16.17	22.9	106.4
5	1993	2939	22.29	26.2	114.7
6	1994	3923	26.87	20.7	124.1
7	1995	4854	33.70	20.3	117.1
8	1996	5576	38.29	17.7	108.3
9	1997	6054	48.60	18.7	102.8
10	1998	6308	50.71	11.9	99.2
11	1999	6551	57.10	14.2	98.6
12	2000	7086	60.70	17.7	100.4
13	2001	7651	70.36	25.6	100.7
14	2002	8214	109.20	30.9	99.2
15	2003	9111	202.01	30.6	101.2
16	2004	10561	231.40	38.9	103.9



图形描述 (例题分析)



统计学
STATISTICS
(第三版)

增长率分析

作者：张占贞，青岛科技大学经济与管理学院

增长率 (growth rate)

1. 也称增长速度
2. 报告期观察值与基期观察值之比减1，用百分比表示
3. 由于对比的基期不同，增长率可以分为环比增长率和定基增长率
4. 由于计算方法的不同，有一般增长率、平均增长率、年度化增长率

环比增长率与定基增长率

1. 环比增长率

- 报告期水平与上一期水平之比减1

$$G_i = \frac{Y_i}{Y_{i-1}} - 1 \quad (i = 1, 2, \dots, n)$$

2. 定基增长率

- 报告期水平与某一固定时期水平之比减1

$$G_i = \frac{Y_i}{Y_0} - 1 \quad (i = 1, 2, \dots, n)$$

平均增长率 (average rate of increase)

1. 序列中各逐期环比值(也称环比发展速度) 的几何平均数减1后的结果
2. 描述现象在整个观察期内平均增长变化的程度
3. 通常用几何平均法求得。计算公式为

$$\begin{aligned}\bar{G} &= \sqrt[n]{\frac{Y_1}{Y_0} \times \frac{Y_2}{Y_1} \times \cdots \times \frac{Y_n}{Y_{n-1}}} - 1 = \sqrt[n]{\prod \frac{Y_i}{Y_{i-1}}} - 1 \\ &= \sqrt[n]{\frac{Y_n}{Y_0}} - 1 \quad (i = 1, 2, \cdots, n)\end{aligned}$$

平均增长率 (例题分析)

【例】见人均GDP数据

年平均增长率为：

$$\bar{G} = \sqrt[n]{\frac{Y_n}{Y_0}} - 1 = \sqrt[14]{\frac{10561}{1634}} - 1 = 114.26\% - 1 = 14.26\%$$

2005年和2006年人均GDP的预测值分别为：

$$\begin{aligned}\hat{Y}_{2005} &= 2004 \text{年数值} \times (1 + \text{年平均增长率}) \\ &= 10561 \times (1 + 14.26\%) = 12067.0\end{aligned}$$

$$\begin{aligned}\hat{Y}_{2006} &= 2004 \text{年数值} \times (1 + \text{年平均增长率})^2 \\ &= 10561 \times (1 + 14.26\%)^2 = 13787.8\end{aligned}$$

增长率分析中应注意的问题

1. 当时间序列中的观察值出现0或负数时，不宜计算增长率
2. 例如：假定某企业连续五年的利润额分别为5，2，0，-3，2万元，对这一序列计算增长率，要么不符合数学公理，要么无法解释其实际意义。在这种情况下，适宜直接用绝对数进行分析
3. 在有些情况下，不能单纯就增长率论增长率，要注意增长率与绝对水平的结合分析

增长率分析中应注意的问题 (例题分析)

【例】假定有两个生产条件基本相同的企业，各年的利润额及有关的速度值如下表

甲、乙两个企业的有关资料				
年 份	甲 企 业		乙 企 业	
	利润额(万元)	增长率(%)	利润额(万元)	增长率(%)
2005	500	—	60	—
2006	600	20	84	40

增长率分析中应注意的问题 (增长1%绝对值)

1. 增长率每增长一个百分点而增加的绝对量
2. 用于弥补增长率分析中的局限性
3. 计算公式为

$$\text{增长1\%绝对值} = \frac{\text{前期水平}}{100}$$

甲企业增长1%绝对值=500/100=5万元

乙企业增长1%绝对值=60/100=0.6万元

§ 13.3 时间序列预测的程序

13.3.1 确定时间序列的成分

13.3.2 选择预测方法

13.3.3 预测方法的评估

统计学
STATISTICS
(第三版)

确定时间序列的成分

作者：张占贞，青岛科技大学经济与管理学院

确定趋势成分 (例题分析)

【例】一种股票连续16周的收盘价如下表所示。试确定其趋势及其类型

	A	B	C	D
1	日期(周)	收盘价格(元)	日期(周)	收盘价格(元)
2	1	15.03	9	9.12
3	2	11.69	10	8.51
4	3	9.63	11	4.45
5	4	10.58	12	4.02
6	5	8.48	13	5.29
7	6	6.98	14	6.51
8	7	6.82	15	6.02
9	8	7.69	16	6.07

确定趋势成分 (例题分析)

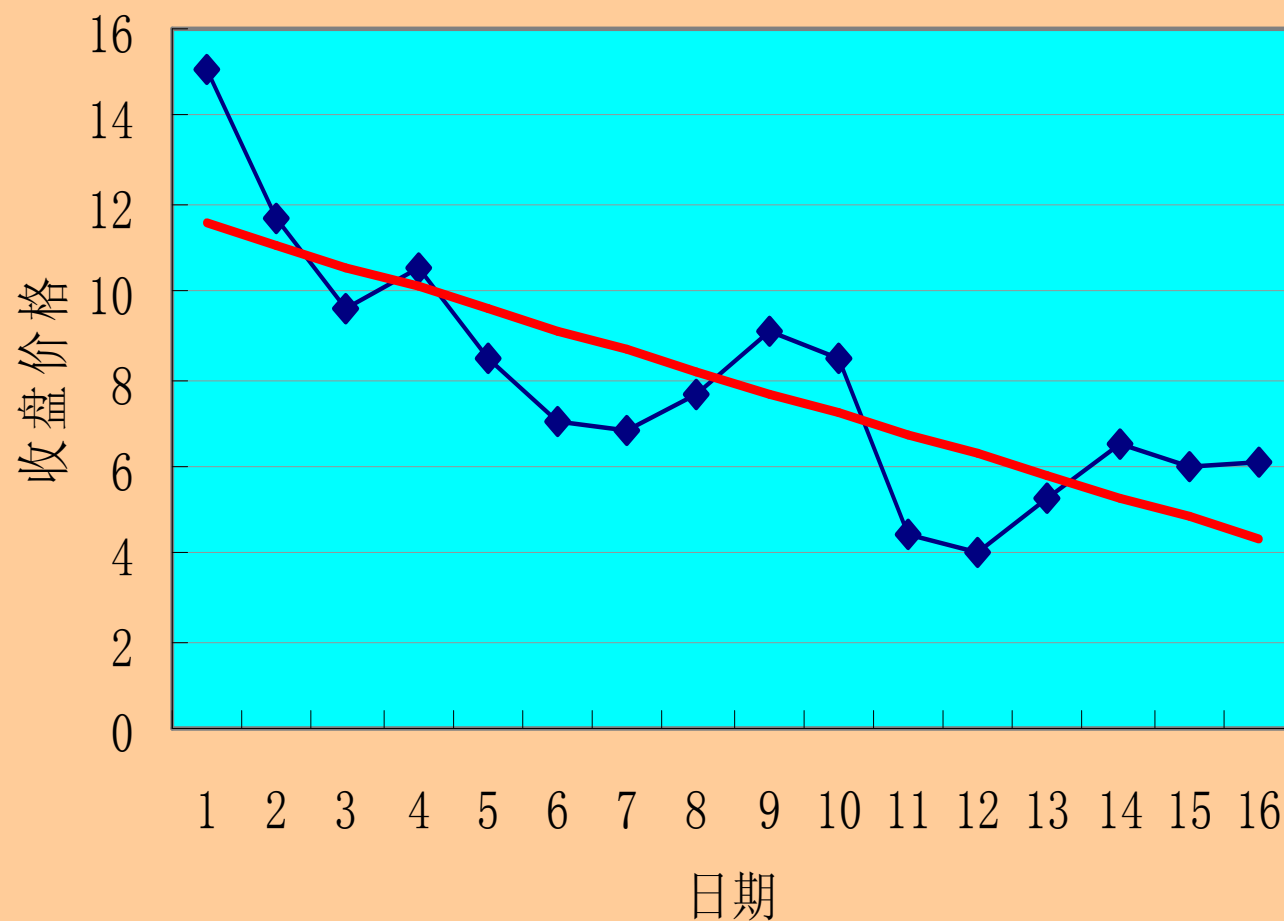
直线趋势方程

$$\hat{Y} = 12.0233 - 0.4815t$$

回归系数检验

$$P=0.000179$$

$$R^2=0.645$$



确定趋势成分 (例题分析)

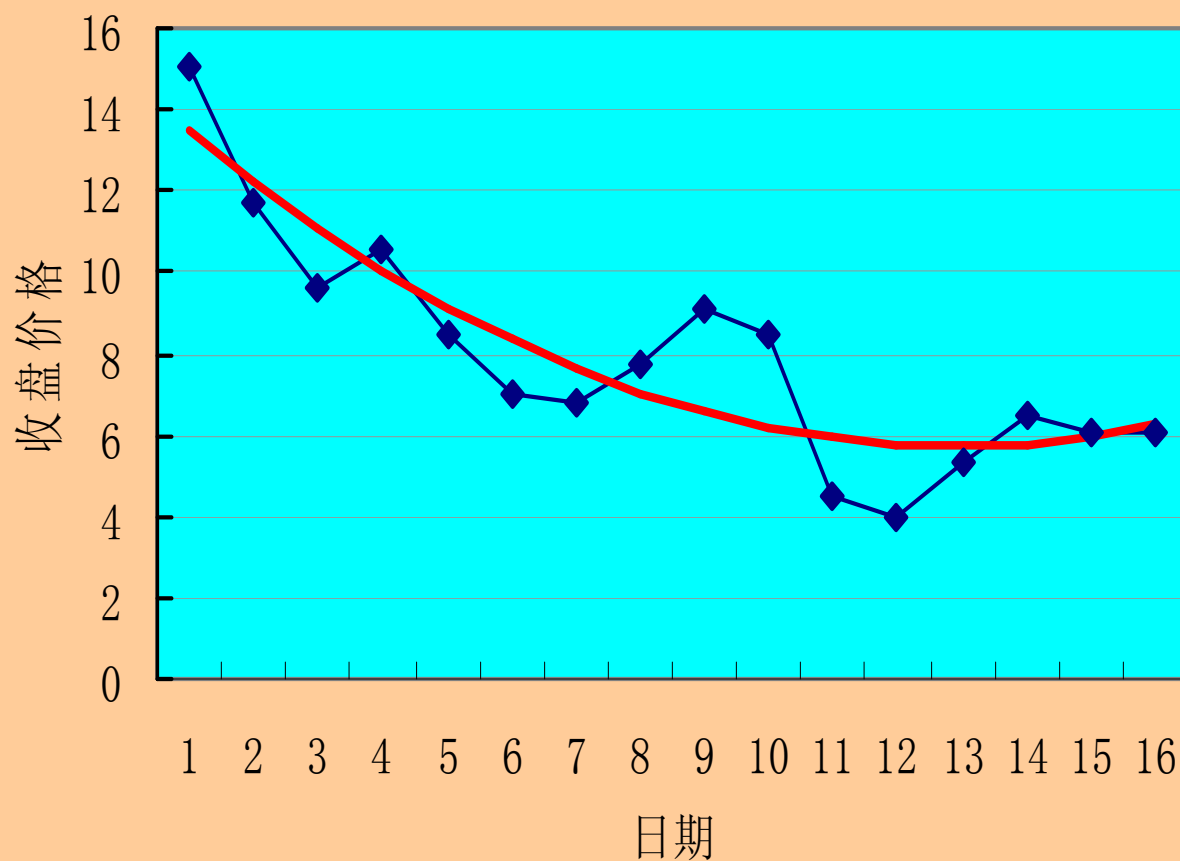
二次曲线方程

$$\hat{Y}=1480541.4088-0.0546x^2$$

回归系数检验

$$P=0.012556$$

$$R^2=0.7841$$



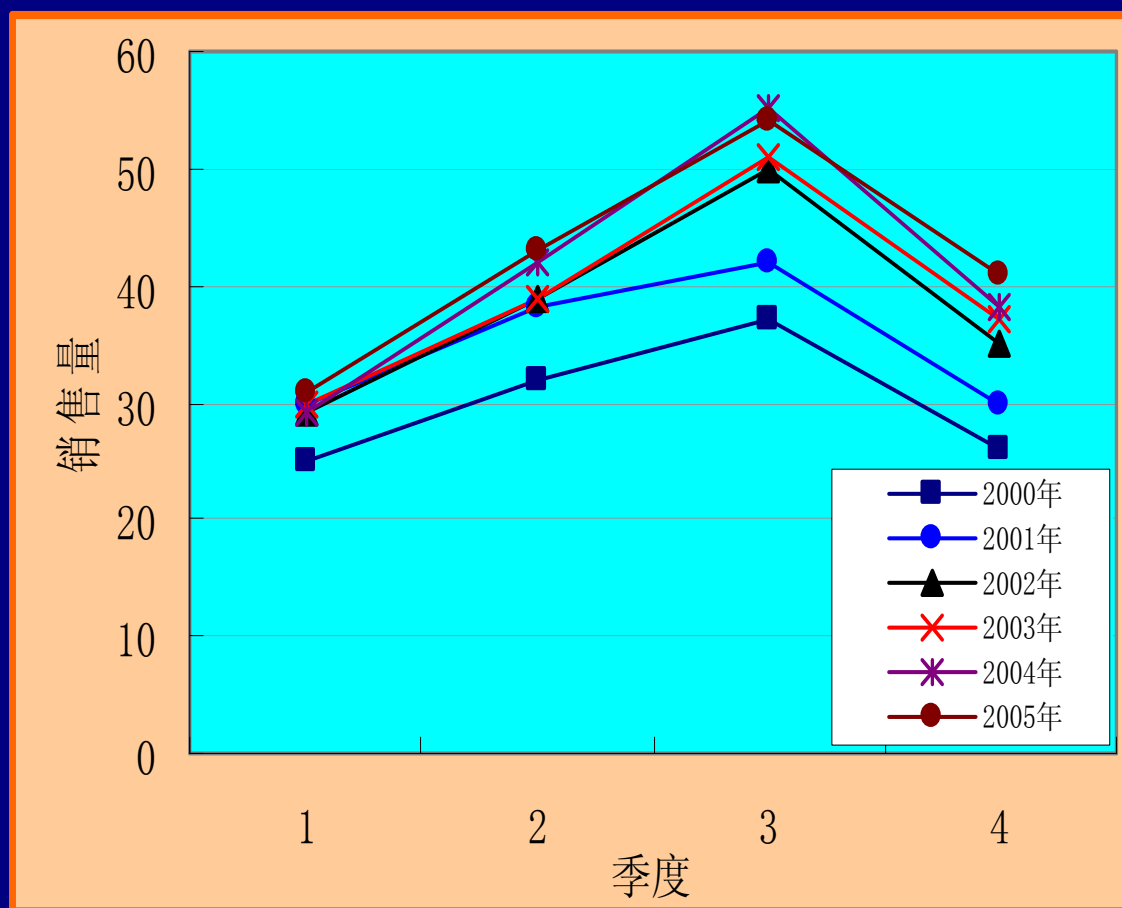
确定季节成分 (例题分析)

【例】下面是一家啤酒生产企业 2000 ~ 2005 年各季度的啤酒销售量数据。试根据这 6 年的数据绘制年度折叠时间序列图，并判断啤酒销售量是否存在季节性

	A	B	C	D	E
1	年份	季 度			
2		1	2	3	4
3	2000	25	32	37	26
4	2001	30	38	42	30
5	2002	29	39	50	35
6	2003	30	39	51	37
7	2004	29	42	55	38
8	2005	31	43	54	41

年度折叠时间序列图 (folded annual time series plot)

1. 将每年的数据分开画在图上
2. 若序列只存在季节成分，年度折叠序列图中的折线将会有交叉
3. 若序列既含有季节成分又含有趋势，则年度折叠时间序列图中的折线将不会有交叉，而且如果趋势是上升的，后面年度的折线将会高于前面年度的折线，如果趋势是下降的，则后面年度的折线将低于前面年度的折线

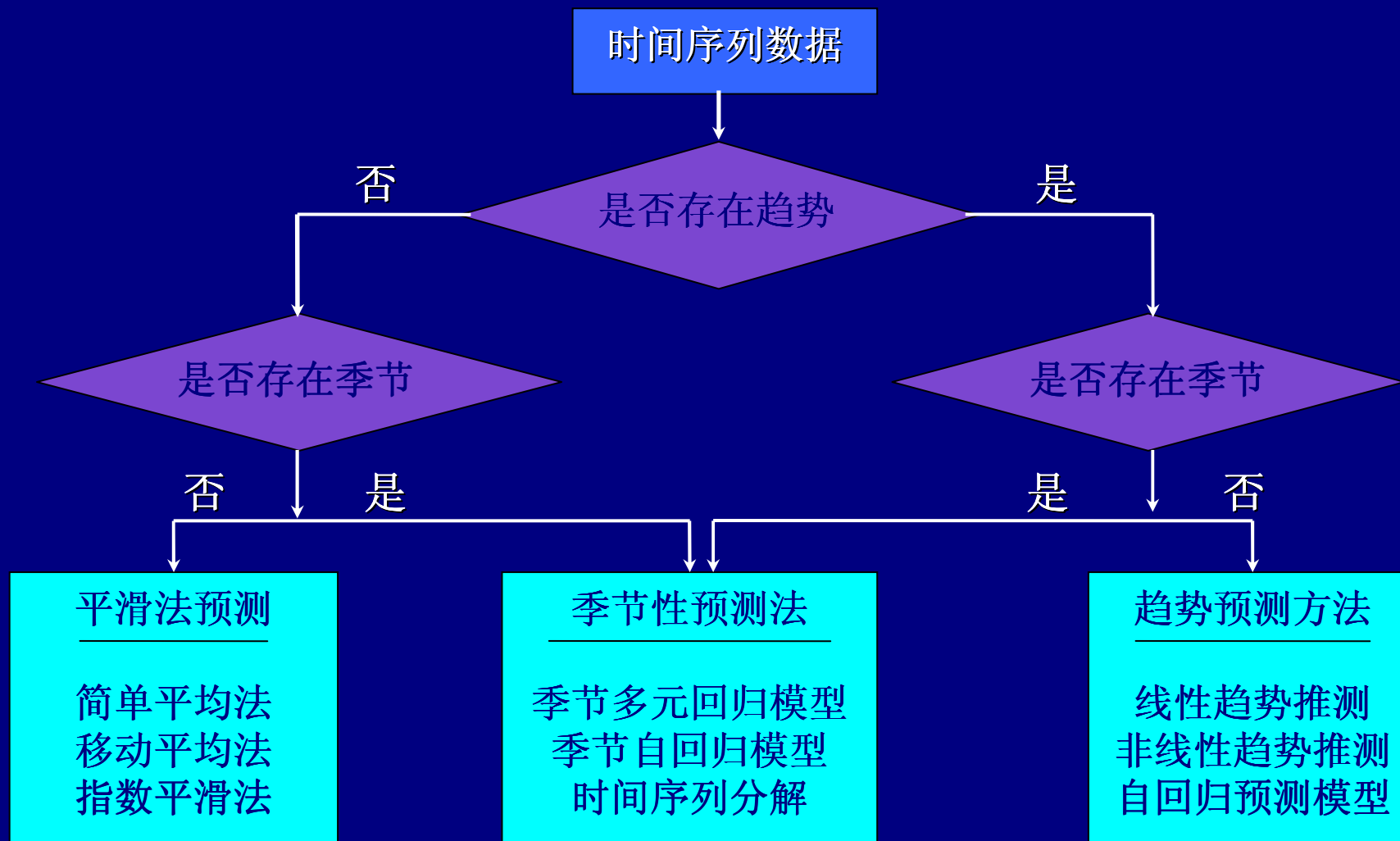


统计学
STATISTICS
(第三版)

选择预测方法

作者：张占贞，青岛科技大学经济与管理学院

预测方法的选择



统计学
STATISTICS
(第三版)

评估预测方法

作者：张占贞，青岛科技大学经济与管理学院

计算误差

1. 平均误差ME(mean error)

$$ME = \frac{\sum_{i=1}^n (Y_i - F_i)}{n}$$

2. 平均绝对误差MAD(mean absolute deviation)

$$MAD = \frac{\sum_{i=1}^n |Y_i - F_i|}{n}$$

计算误差

3、均方误差MSE(mean square error)

$$MSE = \frac{\sum_{i=1}^n (Y_i - F_i)^2}{n}$$

4、平均百分比误差MPE(mean percentage error)

$$MPE = \frac{\sum \left(\frac{Y_i - F_i}{Y_i} \times 100 \right)}{n}$$

5、平均绝对百分比误差MAPE(mean absolute percentage error)

$$MAPE = \frac{\sum_{i=1}^n \left(\frac{|Y_i - F_i|}{Y_i} \times 100 \right)}{n}$$

§ 13.4 平稳序列的预测

13.4.1 简单平均法

13.4.2 移动平均法

13.4.3 指数平滑法

统计学
STATISTICS
(第三版)

简单平均法

作者：张占贞，青岛科技大学经济与管理学院

简单平均法 (simple average)

1. 根据过去已有的 t 期观察值来预测下一期的数值
2. 设时间序列已有的其观察值为 Y_1, Y_2, \dots, Y_t , 则第 $t+1$ 期的预测值 F_{t+1} 为

$$F_{t+1} = \frac{1}{t}(Y_1 + Y_2 + \dots + Y_t) = \frac{1}{t} \sum_{i=1}^t Y_i$$

3. 有了第 $t+1$ 的实际值, 便可计算出预测误差为

$$e_{t+1} = Y_{t+1} - F_{t+1}$$

4. 第 $t+2$ 期的预测值为

$$F_{t+2} = \frac{1}{t+1}(Y_1 + Y_2 + \dots + Y_t + Y_{t+1}) = \frac{1}{t+1} \sum_{i=1}^{t+1} Y_i$$

简单平均法 (特点)

1. 适合对较为平稳的时间序列进行预测
2. 预测结果不准
 - 将远期的数值和近期的数值看作对未来同等重要
 - 从预测角度看，近期的数值要比远期的数值对未来有更大的作用
 - 当时间序列有趋势或有季节变动时，该方法的预测不够准确

统计学
STATISTICS
(第三版)

移动平均法

作者：张占贞，青岛科技大学经济与管理学院

移动平均法 (moving average)

1. 对简单平均法的一种改进方法
2. 通过对时间序列逐期递移求得一系列平均数作为预测值(也可作为趋势值)
3. 有简单移动平均法和加权移动平均法两种

简单移动平均法 (simple moving average)

1. 将最近 k 期数据平均作为下一期的预测值
2. 设移动间隔为 k ($1 < k < t$), 则 t 期的移动平均值为

$$\bar{Y}_t = \frac{Y_{t-k+1} + Y_{t-k+2} + \cdots + Y_{t-1} + Y_t}{k}$$

3. $t+1$ 期的简单移动平均预测值为

$$F_{t+1} = \bar{Y}_t = \frac{Y_{t-k+1} + Y_{t-k+2} + \cdots + Y_{t-1} + Y_t}{k}$$

4. 预测误差用均方误差(MSE) 来衡量

$$MSE = \frac{\text{误差平方和}}{\text{误差个数}} = \frac{\sum_{i=1}^n (Y_i - F_i)^2}{n}$$

简单移动平均法 (特点)

1. 将每个观察值都给予相同的权数
2. 只使用最近期的数据，在每次计算移动平均值时，移动的间隔都为 k
3. 主要适合对较为平稳的序列进行预测
4. 对于同一个时间序列，采用不同的移动步长预测的准确性是不同的
 - 选择移动步长时，可通过试验的办法，选择一个使均方误差达到最小的移动步长

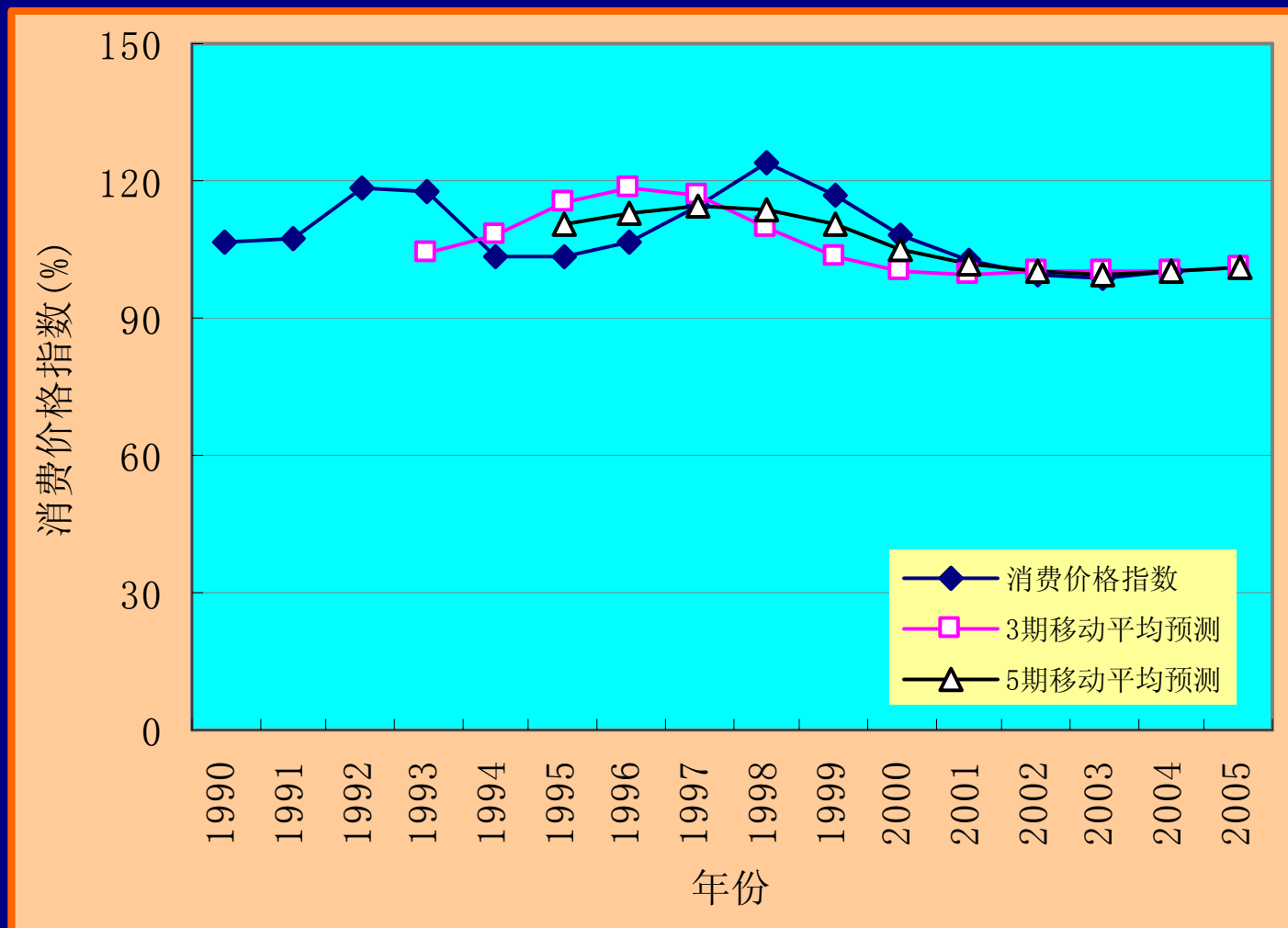
简单移动平均法 (例题分析)

- **【例】**对居民消费价格指数数据，分别取移动间隔 $k=3$ 和 $k=5$ ，用Excel计算各期居民消费价格指数的预测值，计算出预测误差，并将原序列和预测后的序列绘制成图形进行比较

简单移动平均法 (例题分析)

	A	B	C	D	E	F	G	H
1	年份	消费价格 指数	3期移动 平均预测	预测 误差	误差 平方	5期移动 平均预测	预测 误差	误差 平方
2	1990	103.1						
3	1991	103.4						
4	1992	106.4						
5	1993	114.7	104.3	10.4	108.2			
6	1994	124.1	108.2	15.9	253.9			
7	1995	117.1	115.1	2.0	4.1	110.3	6.8	45.7
8	1996	108.3	118.6	-10.3	106.8	113.1	-4.8	23.4
9	1997	102.8	116.5	-13.7	187.7	114.1	-11.3	128.1
10	1998	99.2	109.4	-10.2	104.0	113.4	-14.2	201.6
11	1999	98.6	103.4	-4.8	23.4	110.3	-11.7	136.9
12	2000	100.4	100.2	0.2	0.0	105.2	-4.8	23.0
13	2001	100.7	99.4	1.3	1.7	101.9	-1.2	1.3
14	2002	99.2	99.9	-0.7	0.5	100.3	-1.1	1.3
15	2003	101.2	100.1	1.1	1.2	99.6	1.6	2.5
16	2004	103.9	100.4	3.5	12.5	100.0	3.9	15.1
17	2005		101.4			101.1		
18	合计	—	—	—	803.9		—	579.0

简单移动平均法 (例题分析)



统计学
STATISTICS
(第三版)

指数平滑平均法

作者：张占贞，青岛科技大学经济与管理学院

指数平滑法 (exponential smoothing)

1. 是加权平均的一种特殊形式
2. 对过去的观察值加权平均进行预测的一种方法
3. 观察值时间越远，其权数也跟着呈现指数的下降，因而称为指数平滑
4. 有一次指数平滑、二次指数平滑、三次指数平滑等
5. 一次指数平滑法也可用于对时间序列进行修匀，以消除随机波动，找出序列的变化趋势

一次指数平滑 (single exponential smoothing)

1. 只有一个平滑系数
2. 观察值离预测时期越久远，权数变得越小
3. 以一段时期的预测值与观察值的线性组合作为第 $t+1$ 期的预测值，其预测模型为

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

- Y_t 为第 t 期的实际观察值
- F_t 为第 t 期的预测值
- α 为平滑系数 ($0 < \alpha < 1$)

一次指数平滑

1. 在开始计算时，没有第1期的预测值 F_1 ，通常可以设 F_1 等于第1期的实际观察值，即 $F_1=Y_1$

2. 第2期的预测值为

$$F_2 = \alpha Y_1 + (1-\alpha)F_1 = \alpha Y_1 + (1-\alpha)Y_1 = Y_1$$

3. 第3期的预测值为

$$F_3 = \alpha Y_2 + (1-\alpha)F_2 = \alpha Y_2 + (1-\alpha)Y_1$$

一次指数平滑 (预测误差)

1. 预测精度，用误差均方来衡量

$$\begin{aligned}F_{t+1} &= \alpha Y_t + (1 - \alpha) F_t \\&= \alpha Y_t + F_t - \alpha F_t \\&= F_t + \alpha(Y_t - F_t)\end{aligned}$$

2. F_{t+1} 是第 t 期的预测值 F_t 加上用 α 调整的第 t 期的预测误差 $(Y_t - F_t)$

一次指数平滑 (α 的确定)

1. 不同的 α 会对预测结果产生不同的影响

- 当时间序列有较大的随机波动时，宜选较大的 α ，以便能很快跟上近期的变化
- 当时间序列比较平稳时，宜选较小的 α

2. 选择 α 时，还应考虑预测误差

- 误差均方来衡量预测误差的大小
- 确定 α 时，可选择几个进行预测，然后找出预测误差最小的作为最后的值

一次指数平滑 (例题分析)

【例】对居民消费价格指数数据，选择适当的平滑系数 α ，采用Excel进行指数平滑预测，计算出预测误差，并将原序列和预测后的序列绘制成图形进行比较

第1步：选择【工具】下拉菜单

第2步：选择【数据分析】，并选择【指数平滑】，然后【确定】

第3步：当对话框出现时

在【输入区域】中输入数据区域

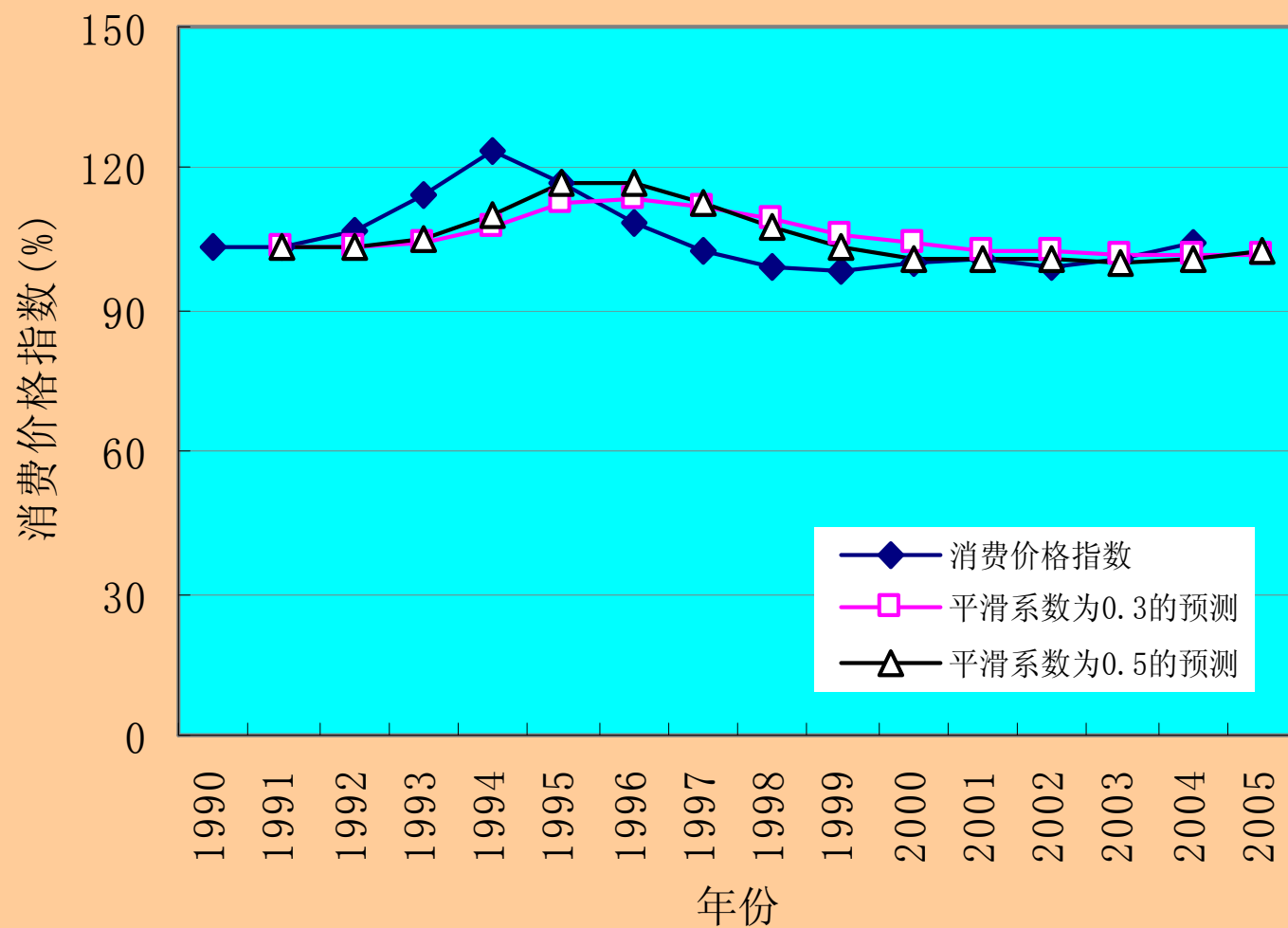
在【阻尼系数】(注意：阻尼系数= $1 - \alpha$)输入的值

选择【确定”】

一次指数平滑 (例题分析)

	A	B	C	D	E	F
1	年份	消费价格指数	$\alpha = 0.3$	误差平方	$\alpha = 0.5$	误差平方
2	1990	103.1				
3	1991	103.4	103.1	0.1	103.1	0.1
4	1992	106.4	103.2	10.3	103.3	9.9
5	1993	114.7	104.2	111.2	104.8	97.5
6	1994	124.1	107.3	281.7	109.8	205.6
7	1995	117.1	112.4	22.5	116.9	0.0
8	1996	108.3	113.8	30.0	117.0	76.0
9	1997	102.8	112.1	87.1	112.7	97.2
10	1998	99.2	109.3	102.7	107.7	72.7
11	1999	98.6	106.3	59.2	103.5	23.7
12	2000	100.4	104.0	12.9	101.0	0.4
13	2001	100.7	102.9	4.9	100.7	0.0
14	2002	99.2	102.2	9.3	100.7	2.3
15	2003	101.2	101.3	0.0	100.0	1.6
16	2004	103.9	101.3	6.8	100.6	11.0
17	2005		102.1		102.2	
18	合计	—	—	738.7	—	597.9

一次指数平滑 (例题分析)



§ 13.5 趋势型序列的预测

13.5.1 线性趋势预测

13.5.2 非线性趋势预测

趋势序列及其预测方法

1. 趋势(trend)

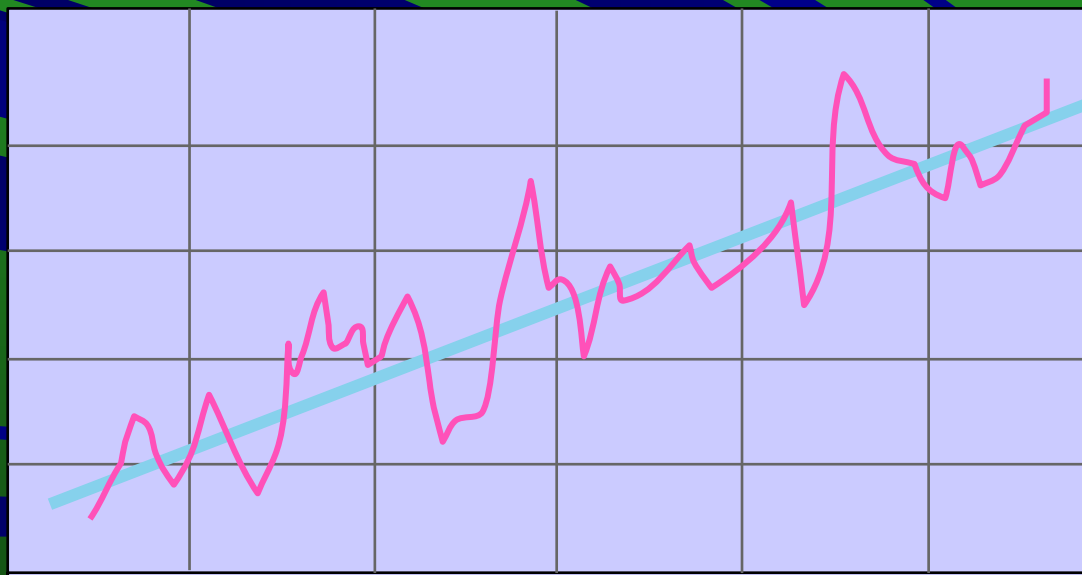
- 持续向上或持续下降的状态或规律

2. 有线性趋势和非线性趋势

3. 方法主要有

- 线性趋势预测
- 非线性趋势预测
- 自回归模型预测

线性趋势预测



线性趋势 (linear trend)



1. 现象随着时间的推移而呈现出稳定增长或下降的线性变化规律
2. 由影响时间序列的基本因素作用形成
3. 时间序列的成分之一
4. 预测方法：线性模型法

线性模型法 (线性趋势方程)

■ ➡ 线性方程的形式为

$$\hat{Y}_t = b_0 + b_1 t$$

- \hat{Y}_t —时间序列的预测值
- t —时间标号
- b_0 —趋势线在 Y 轴上的截距
- b_1 —趋势线的斜率，表示时间 t 变动一个单位时观察值的平均变动数量

线性模型法 (a 和 b 的求解方程)

1. 根据最小二乘法得到求解 b_0 和 b_1 的标准方程为

$$\begin{cases} \sum Y = nb_0 + b_1 \sum t \\ \sum tY = b_0 \sum t + b_1 \sum t^2 \end{cases} \quad \text{解得} \quad \begin{cases} b_1 = \frac{n \sum tY - \sum t \sum Y}{n \sum t^2 - (\sum t)^2} \\ b_0 = \bar{Y} - b_1 \bar{t} \end{cases}$$

2. 预测误差可用估计标准误差来衡量

$$s_Y = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n - m}}$$

m 为趋势方程中待确定的未知常数的个数

线性模型法 (例题分析)

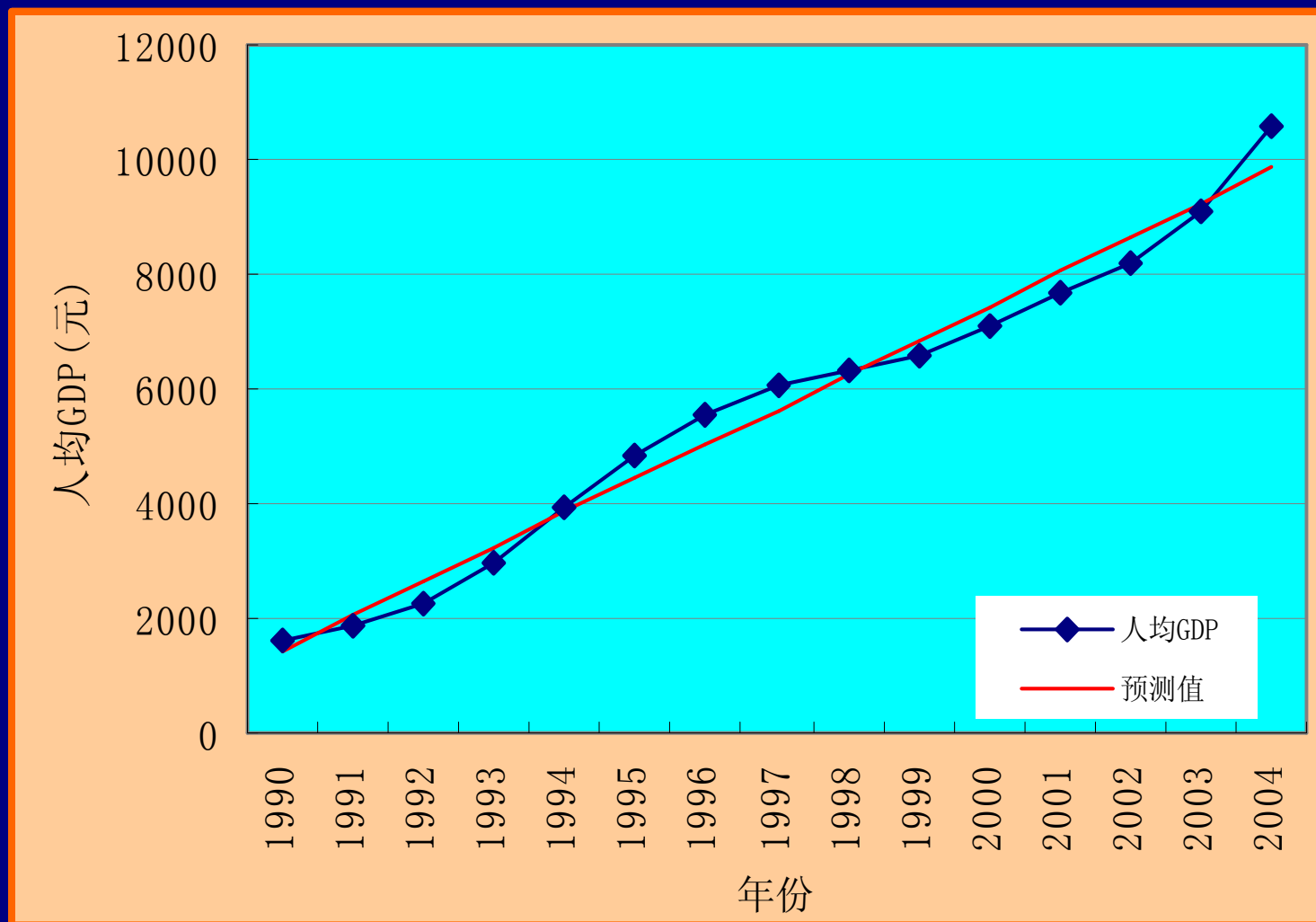
【例】根据人均GDP数据，根据最小二乘法确定直线趋势方程，计算出各期的预测值和预测误差，预测2005年的人均GDP，并将原序列和各期的预测值序列绘制成图形进行比较

1. 线性趋势方程: $\hat{Y}_t = 842.45 + 599.93t$
2. 预测的 R^2 和估计标准误差: $R^2=0.9806$ $s_Y = 932.06$
3. 2005年人口自然增长率的预测值

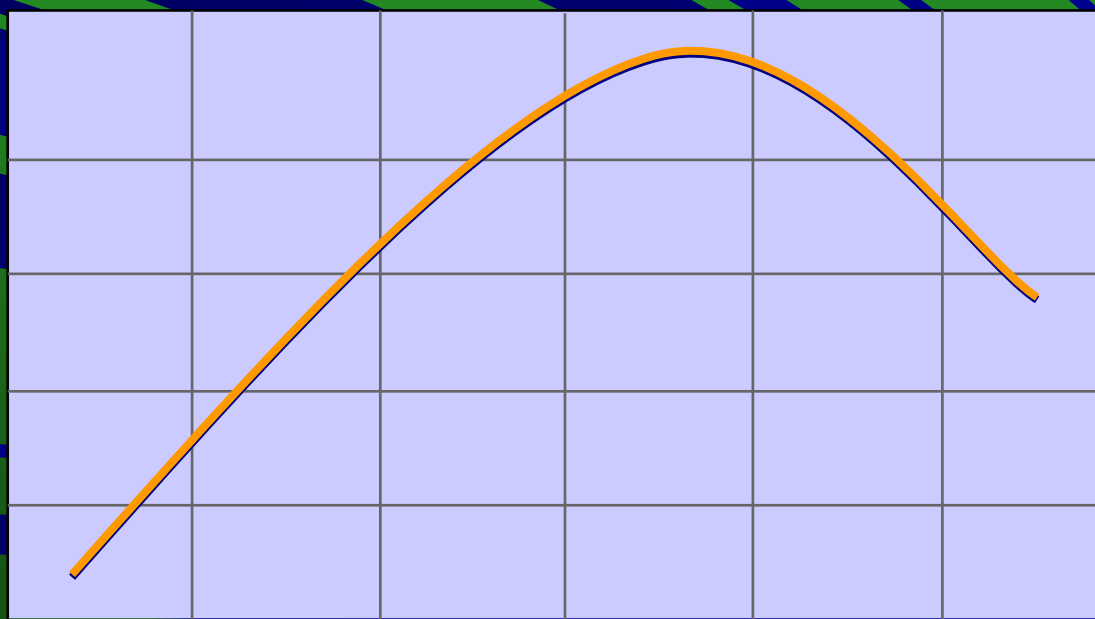
$$\hat{Y}_{2005} = 842.45 + 599.93 \times 16 = 10441.33$$



线性模型法 (例题分析)



非线性趋势预测



指数曲线 (exponential curve)

1. 时间序列以几何级数递增或递减
2. 一般形式为

$$\hat{Y}_t = b_0 b_1^t$$

- b_0, b_1 为待定系数
- 若 $b_1 > 1$, 增长率随着时间 t 的增加而增加
- 若 $b_1 < 1$, 增长率随着时间 t 的增加而降低
- 若 $b_0 > 0, b_1 < 1$, 趋势值逐渐降低到以 0 为极限

指数曲线 (a, b 的求解方法)

1. 采取“线性化”手段将其化为对数直线形式
2. 根据最小二乘法，得到求解 $\lg b_0$ 、 $\lg b_1$ 的标准方程为

$$\begin{cases} \sum \lg Y = n \lg b_0 + \lg b_1 \sum t \\ \sum t \lg Y = \lg b_0 \sum t + \lg b_1 \sum t^2 \end{cases}$$

3. 求出 $\lg b_0$ 和 $\lg b_1$ 后，再取其反对数，即得算术形式的 b_0 和 b_1

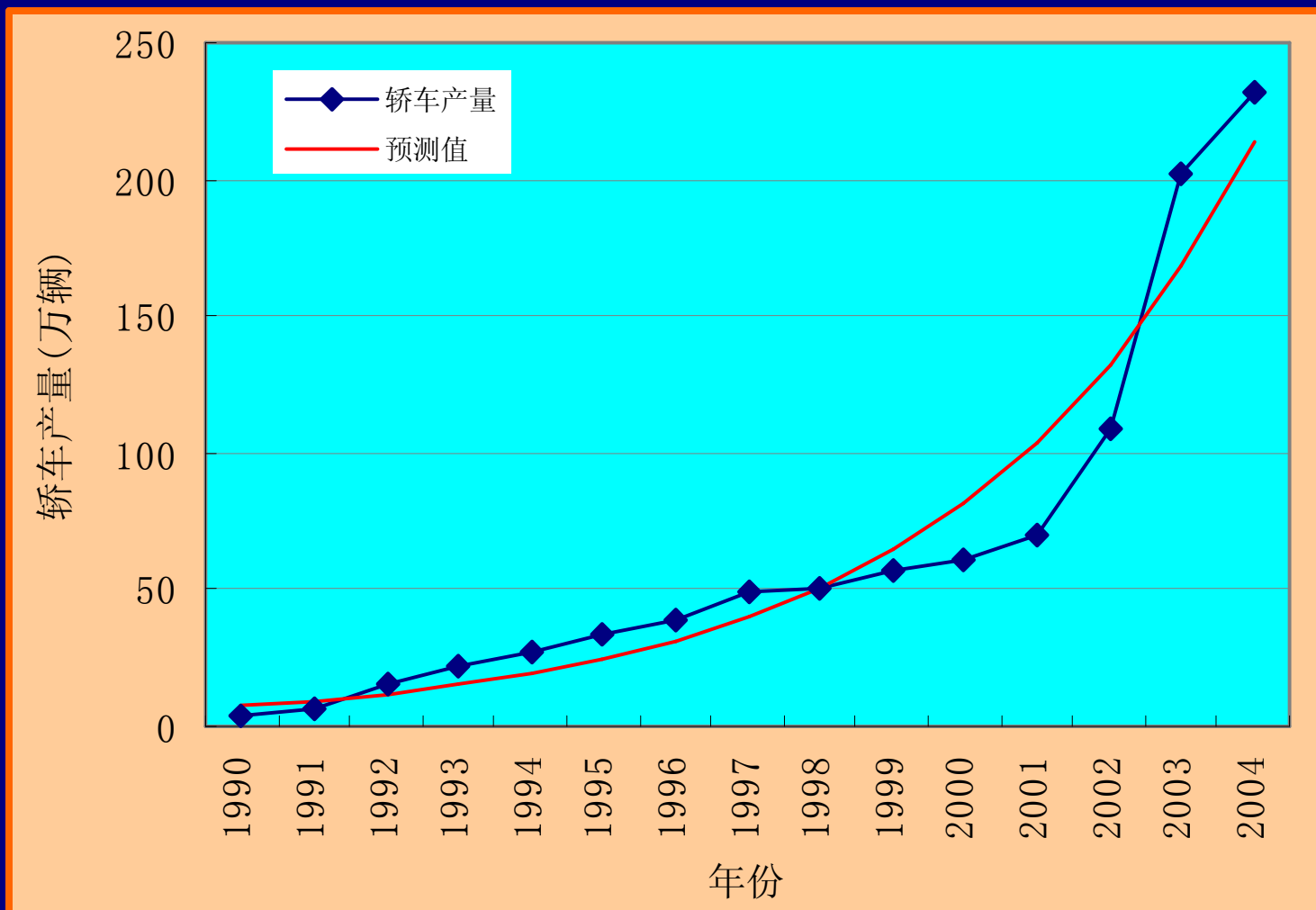
指数曲线 (例题分析)

【例】根据轿车产量数据，确定指数曲线方程，计算出各期的预测值和预测误差，预测2005年的轿车产量，并将原序列和各期的预测值序列绘制成图形进行比较

1. 指数曲线趋势方程： $\hat{Y}_t = 5.74637 \times (1.27286)^t$
2. 预测的估计标准误差： $s_Y = 303$
3. 2005年轿车产量的预测值

$$\hat{Y}_{2005} = 5.74637 \times (1.27286)^{16} = 272.82$$

指数曲线 (例题分析)



指数曲线与直线的比较

1. 比一般的趋势直线有着更广泛的应用
2. 可以反应现象的相对发展变化程度
 - 上例中， $b_1=1.27286$ 表示1990—2004年轿车产量的年平均增长率为27.286%
3. 不同序列的指数曲线可以进行比较
 - 比较分析相对增长程度

修正指数曲线 (modified exponential curve)

1. 在一般指数曲线的方程上增加一个常数项 K
2. 一般形式为

$$\hat{Y}_t = K + b_0 b_1^t$$

- K, b_0, b_1 为待定系数
- $K > 0, b_0 \neq 0, 0 < b_1 \neq 1$

3. 用于描述的现象：初期增长迅速，随后增长率逐渐降低，最终则以 K 为增长极限

修正指数曲线 (求解 k , b_0 , b_1 的三和法)

1. 趋势值 K 无法事先确定时采用
2. 将时间序列观察值等分为三个部分，每部分有 m 个时期
3. 令预测值的三个局部总和分别等于原序列观察值的三个局部总和

修正指数曲线 (求解 k , b_0 , b_1 的三和法)

1. 设观察值的三个局部总和分别为 S_1 , S_2 , S_3

$$S_1 = \sum_{t=1}^m Y_t, \quad S_2 = \sum_{t=m+1}^{2m} Y_t, \quad S_3 = \sum_{t=2m+1}^{3m} Y_t$$

2. 根据三和法求得

$$\begin{cases} b_1 = \left(\frac{S_3 - S_2}{S_2 - S_1} \right)^{\frac{1}{m}} \\ b_0 = (S_2 - S_1) \frac{b_1 - 1}{b_1 (b_1^m - 1)^2} \\ K = \frac{1}{m} \left(S_1 - \frac{b_0 b_1 (b_1^m - 1)}{b_1 - 1} \right) \end{cases}$$

作者：张占贞，青岛科技大学经济与管理学院

修正指数曲线 (例题分析)

【例】我国1990—2004年城镇新建住宅面积数据如右表所示。试确定修正指数曲线方程，计算出各期的预测值和预测误差，预测2005年的城镇新建住宅面积，并将原序列和各期的预测值序列绘制成图形进行比较

	A	B	C	D
1	年 份	城镇新建住宅面积 (亿平方米)	年 份	城镇新建住宅面积 (亿平方米)
2	1990	1.73	1998	4.76
3	1991	1.92	1999	5.59
4	1992	2.40	2000	5.49
5	1993	3.08	2001	5.75
6	1994	3.57	2002	5.98
7	1995	3.75	2003	5.50
8	1996	3.95	2004	5.69
9	1997	4.06		

修正指数曲线 (例题分析)

	A	B	C	D	E	F
1	年份	t	新建住宅面积	预测值	残差	残差平方
2	1990	1	1.73	1.59	0.14	0.02
3	1991	2	1.92	2.10	-0.18	0.03
4	1992	3	2.40	2.58	-0.18	0.03
5	1993	4	3.08	3.01	0.07	0.00
6	1994	5	3.57	3.42	0.15	0.02
7	S1	—	12.70	12.70	—	0.11
8	1995	6	3.75	3.79	-0.04	0.00
9	1996	7	3.95	4.13	-0.18	0.03
10	1997	8	4.06	4.45	-0.39	0.15
11	1998	9	4.76	4.74	0.02	0.00
12	1999	10	5.59	5.01	0.58	0.34
13	S2	—	22.11	22.11	—	0.52
14	2000	11	5.49	5.26	0.23	0.05
15	2001	12	5.75	5.49	0.26	0.07
16	2002	13	5.98	5.70	0.28	0.08
17	2003	14	5.50	5.89	-0.39	0.16
18	2004	15	5.69	6.07	-0.38	0.15
19	S3	—	28.41	28.41	—	0.51

修正指数曲线 (例题分析)

解得 K , b_0 , b_1 如下

$$\begin{cases} b = \left(\frac{28.41 - 22.11}{22.11 - 12.70} \right)^{\frac{1}{5}} = 0.9229 \\ a = (22.11 - 12.70) \times \frac{0.9229 - 1}{0.9229 \times (0.9229^5 - 1)^2} = -7.1979 \\ K = \frac{1}{5} \times \left(12.70 - \frac{-7.1979 \times 0.9229 \times (0.9229^5 - 1)}{0.9229 - 1} \right) = 8.2344 \end{cases}$$

修正指数曲线 (例题分析)

➡新建住宅面积的修正指数曲线方程

$$\hat{Y}_t = 8.2344 - 7.1979 \times 0.9229^t$$

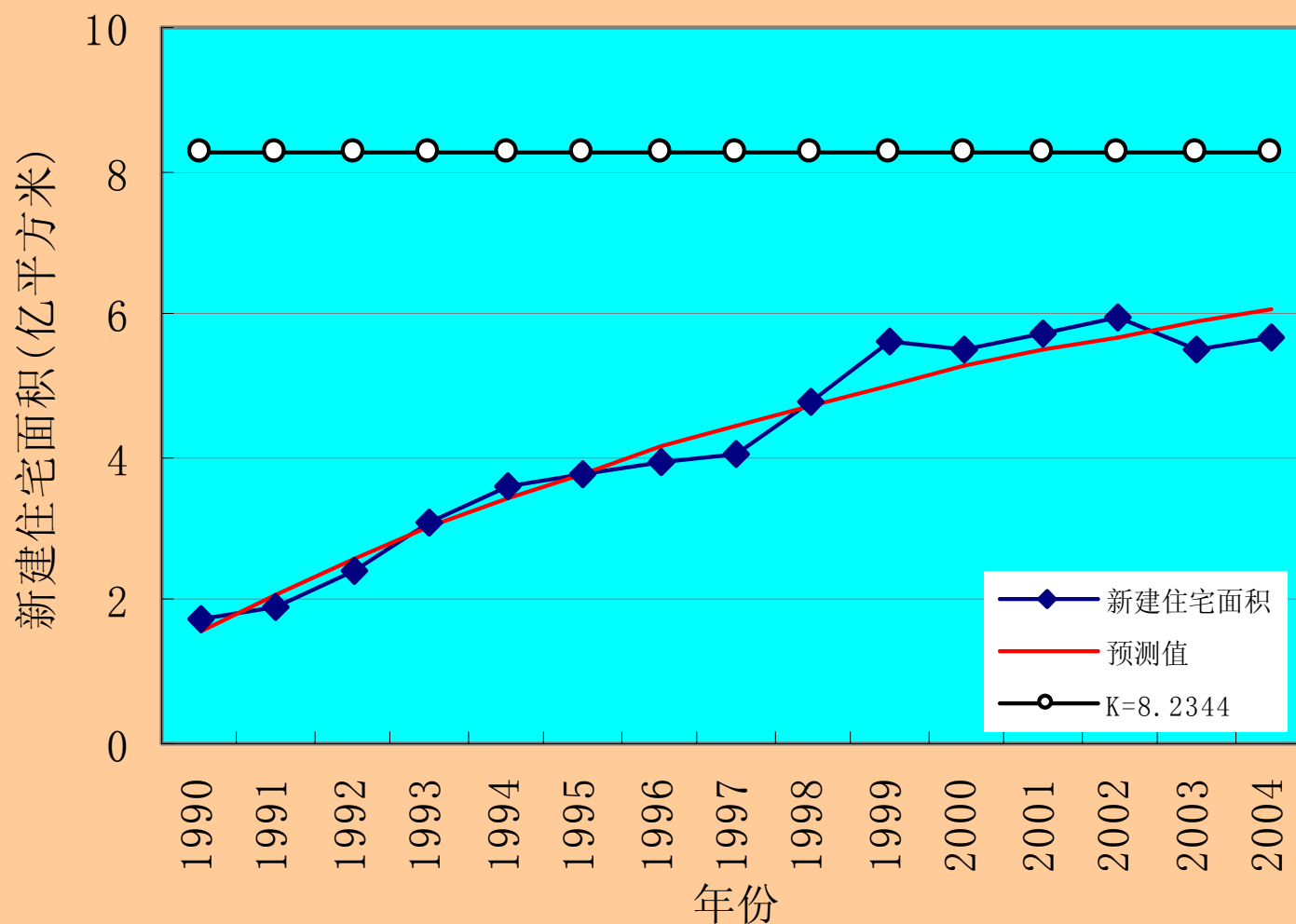
➡2005年的预测值

$$\hat{Y}_t = 8.2344 - 7.1979 \times 0.9229^{16} = 6.24$$

➡预测的估计标准误差

$$s_Y = 0.31$$

修正指数曲线 (例题分析)



Gompertz 曲线 (Gompertz curve)

1. 以英国统计学家和数学家 B.Gompertz 的名字而命名
2. 一般形式为

$$\hat{Y}_t = Kb_0^{b_1^t}$$

- K, b_0, b_1 为待定系数
 - $K > 0, 0 < b_0 \neq 1, 0 < b_1 \neq 1$
3. 描述的现象：初期增长缓慢，以后逐渐加快，当达到一定程度后，增长率又逐渐下降，最后接近一条水平线
 4. 两端都有渐近线，上渐近线为 $Y \rightarrow K$, 下渐近线为 $Y \rightarrow 0$

Gompertz 曲线

(求解 k , b_0 , b_1 的三和法)

1. 将其改写为对数形式: $\lg \hat{Y}_t = \lg K + (\lg b_0)b_1^t$
2. 仿照修正指数曲线的常数确定方法, 求出 $\lg b_0$ 、 $\lg K$ 、 b_1
3. 取 $\lg b_0$ 、 $\lg K$ 的反对数求得 b_0 和 K

$$\text{令: } S_1 = \sum_{t=1}^m \lg Y_t, \quad S_2 = \sum_{t=m+1}^{2m} \lg Y_t, \quad S_3 = \sum_{t=2m+1}^{3m} \lg Y_t$$

$$\text{则有: } \begin{cases} b_1 = \left(\frac{S_3 - S_2}{S_2 - S_1} \right)^{\frac{1}{m}} \\ \lg b_0 = (S_2 - S_1) \frac{b_1 - 1}{b_1 (b_1^m - 1)^2} \\ \lg K = \frac{1}{m} \left(S_1 - \frac{b_1 (b_1^m - 1)}{b_1 - 1} \bullet \lg b_0 \right) \end{cases}$$

Gompertz 曲线 (例题分析)

【例】我国1990—2004年城镇新建住宅面积数据如右表所示。试确定修正指数曲线方程，计算出各期的预测值和预测误差，预测2005年的城镇新建住宅面积，并将原序列和各期的预测值序列绘制成图形进行比较

	A	B	C	D
1	年 份	城镇新建住宅面积 (亿平方米)	年 份	城镇新建住宅面积 (亿平方米)
2	1990	1.73	1998	4.76
3	1991	1.92	1999	5.59
4	1992	2.40	2000	5.49
5	1993	3.08	2001	5.75
6	1994	3.57	2002	5.98
7	1995	3.75	2003	5.50
8	1996	3.95	2004	5.69
9	1997	4.06		

Gompertz 曲线

(例题分析)

	A	B	C	D	E	F	G
1	年份	t	新建住宅面积	Lg(Y)	预测值	残差	残差平方
2	1990	1	1.73	0.238046	1.70	0.03	0.00
3	1991	2	1.92	0.283301	2.10	-0.18	0.03
4	1992	3	2.40	0.380211	2.51	-0.11	0.01
5	1993	4	3.08	0.488551	2.92	0.16	0.02
6	1994	5	3.57	0.552668	3.33	0.24	0.06
7	S1	—	12.70	1.942778	12.57		0.13
8	1995	6	3.75	0.574031	3.72	0.03	0.00
9	1996	7	3.95	0.596597	4.08	-0.13	0.02
10	1997	8	4.06	0.608526	4.43	-0.37	0.13
11	1998	9	4.76	0.677607	4.74	0.02	0.00
12	1999	10	5.59	0.747412	5.02	0.57	0.32
13	S2	—	22.11	3.204173	21.99		0.47
14	2000	11	5.49	0.739572	5.28	0.21	0.04
15	2001	12	5.75	0.759668	5.51	0.24	0.06
16	2002	13	5.98	0.776701	5.71	0.27	0.07
17	2003	14	5.50	0.740363	5.89	-0.39	0.15
18	2004	15	5.69	0.755112	6.04	-0.35	0.13
19	S3	—	28.41	3.771416	28.43		0.45

Gompertz 曲线 (例题分析)

Gompertz 曲线计算过程

$$\begin{cases} b_1 = \left(\frac{3.771416 - 3.204173}{3.204173 - 1.942778} \right)^{\frac{1}{5}} = 0.852283 \\ \lg b_0 = (3.204173 - 1.942778) \times \frac{0.852283 - 1}{0.852283 \times (0.852283^5 - 1)^2} = -0.721926 \\ \lg K = \frac{1}{5} \left(1.942778 - \frac{0.852283 \times (0.852283^5 - 1)}{0.852283 - 1} \times (-0.721926) \right) = 0.846991 \end{cases}$$

Gompertz 曲线 (例题分析)

➡ 新建住宅面积的Gompertz曲线方程

$$\hat{Y}_t = 0.852283 \times 0.189703^{7.030578^t}$$

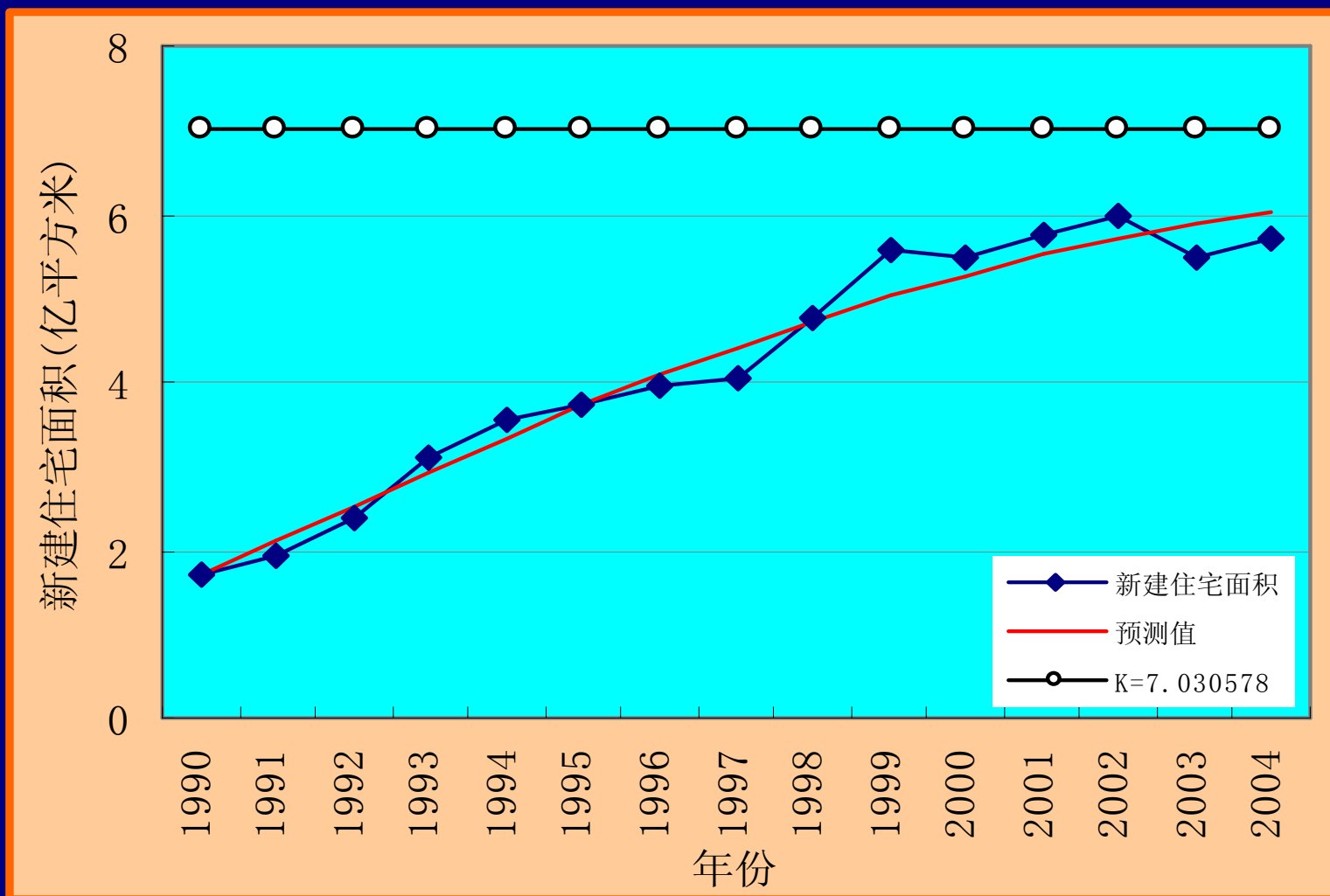
➡ 2005年的预测值

$$\hat{Y}_t = 0.852283 \times 0.189703^{7.030578^{16}} = 6.18$$

➡ 预测的估计标准误差

$$s_Y = 0.30$$

Gompertz 曲线 (例题分析)



多阶曲线

1. 有些现象的变化形态比较复杂，它们不是按照某种固定的形态变化，而是有升有降，在变化过程中可能有几个拐点。这时就需要拟合多项式函数
2. 当只有一个拐点时，可以拟合二阶曲线，即抛物线；当有两个拐点时，需要拟合三阶曲线；当有 $k-1$ 个拐点时，需要拟合 k 阶曲线
3. k 阶曲线函数的一般形式为

$$\hat{Y}_t = b_0 + b_1 t + b_2 t^2 + \cdots + b_k t^k$$

4. 线性化后，根据最小二乘法求 $b_0, b_1, b_2, \cdots, b_k$

多阶曲线 (例题分析)

【例】根据的金属切削机床产量数据，拟合适当的趋势曲线，计算出各期的预测值和预测误差，预测2005年的金属切削机床产量，并将原序列和各期的预测值序列绘制成图形进行比较

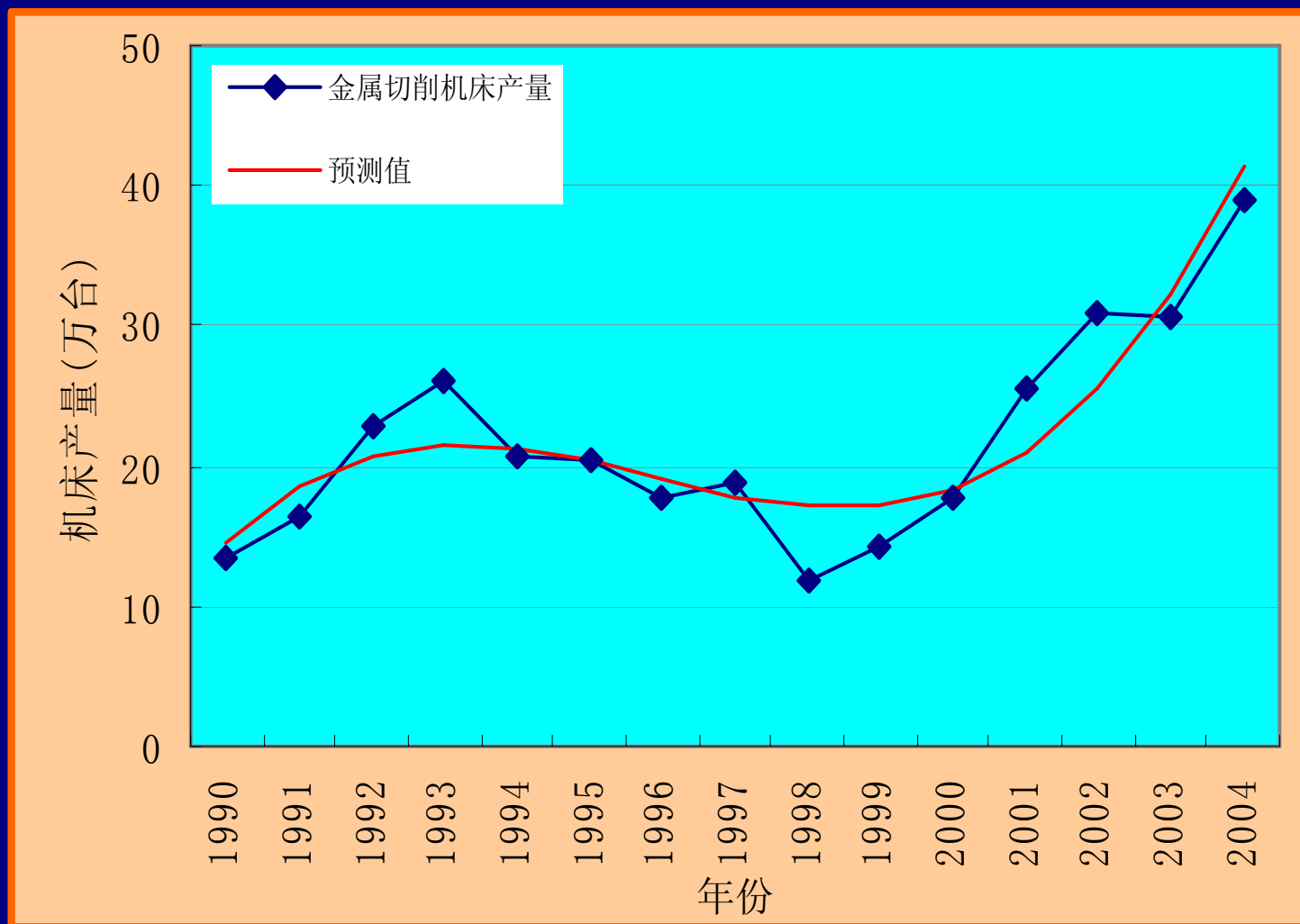
1. 三阶曲线方程: $\hat{Y}_t = 8.5360 + 7.2507t - 1.2448t^2 + 0.0605t^3$

2. 2005年的预测值

$$\hat{Y}_{2005} = 8.5360 + 7.2507 \times 16 - 1.2448 \times 16^2 + 0.0605 \times 16^3 = 53.69$$

3. 预测的估计标准误差: $s_Y = 3.33$

多阶曲线 (例题分析)



趋势线的选择

1. 观察散点图
2. 根据观察数据本身，按以下标准选择趋势线
 - 一次差大体相同，配合直线
 - 二次差大体相同，配合二次曲线
 - 对数的一次差大体相同，配合指数曲线
 - 一次差的环比值大体相同，配合修正指数曲线
 - 对数一次差的环比值大体相同，配合Gompertz曲线
 - 倒数一次差的环比值大体相同，配合Logistic曲线
3. 比较估计标准误差

§ 13.6 复合型序列的分解预测

13.6.1 确定并分离季节成分

13.6.2 建立预测模型并进行预测

13.6.3 计算最后的预测值

预测步骤

1. 确定并分离季节成分

- 计算季节指数，以确定时间序列中的季节成分
- 将季节成分从时间序列中分离出去，即用每一个观测值除以相应的季节指数，以消除季节性

2. 建立预测模型并进行预测

- 对消除季节成分的序列建立适当的预测模型，并根据这一模型进行预测

3. 计算出最后的预测值

- 用预测值乘以相应的季节指数，得到最终的预测值

统计学
STATISTICS
(第三版)

确定并分离季节成分

作者：张占贞，青岛科技大学经济与管理学院

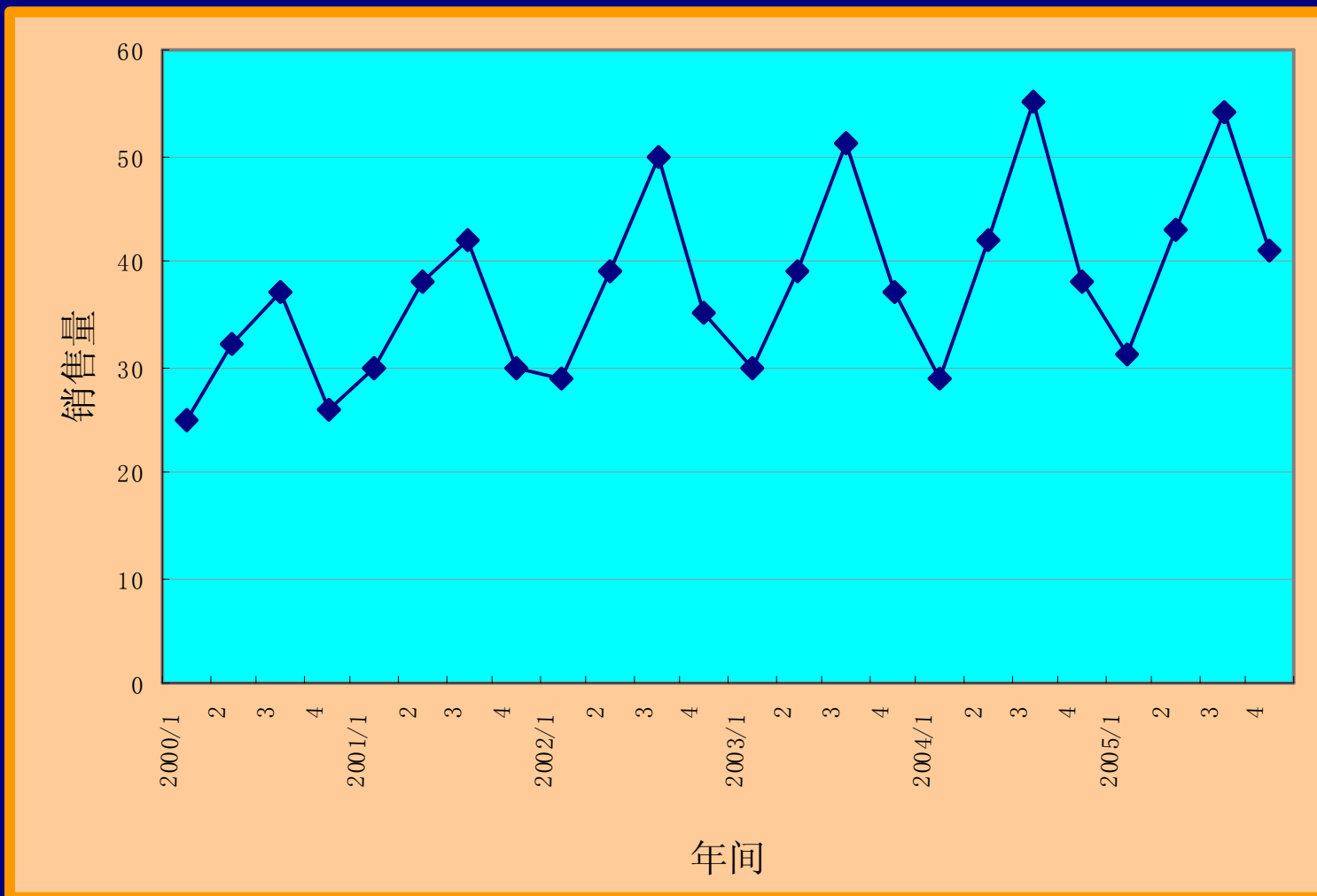
季节指数 (例题分析)

■ 【例】下表是一家啤酒生产企业2000—2005年各季度的啤酒销售量数据。试计算各季的季节指数。

	A	B	C	D	E
1	年份	季 度			
2		1	2	3	4
3	2000	25	32	37	26
4	2001	30	38	42	30
5	2002	29	39	50	35
6	2003	30	39	51	37
7	2004	29	42	55	38
8	2005	31	43	54	41



图形描述



计算季节指数 (seasonal index)

1. 刻画序列在一个年度内各月或季的典型季节特征
2. 以其平均数等于100%为条件而构成
3. 反映某一月份或季度的数值占全年平均数值的大小
4. 如果现象的发展没有季节变动，则各期的季节指数应等于100%
5. 季节变动的程度是根据各季节指数与其平均数(100%)的偏差程度来测定
 - 如果某一月份或季度有明显的季节变化，则各期的季节指数应大于或小于100%

季节指数 (计算步骤)

1. 计算移动平均值(季度数据采用4项移动平均, 月份数据采用12项移动平均), 并将其结果进行“中心化”处理
 - 将移动平均的结果再进行一次2项的移动平均, 即得出“中心化移动平均值”(CMA)
2. 计算移动平均的比值, 也称为季节比率
 - 将序列的各观察值除以相应的中心化移动平均值, 然后再计算出各比值的季度(或月份)平均值, 即季节指数
3. 季节指数调整
 - 各季节指数的平均数应等于1或100%, 若根据第2步计算的季节比率的平均值不等于1时, 则需要进行调整
 - 具体方法是: 将第2步计算的每个季节比率的平均值除以它们的总平均值

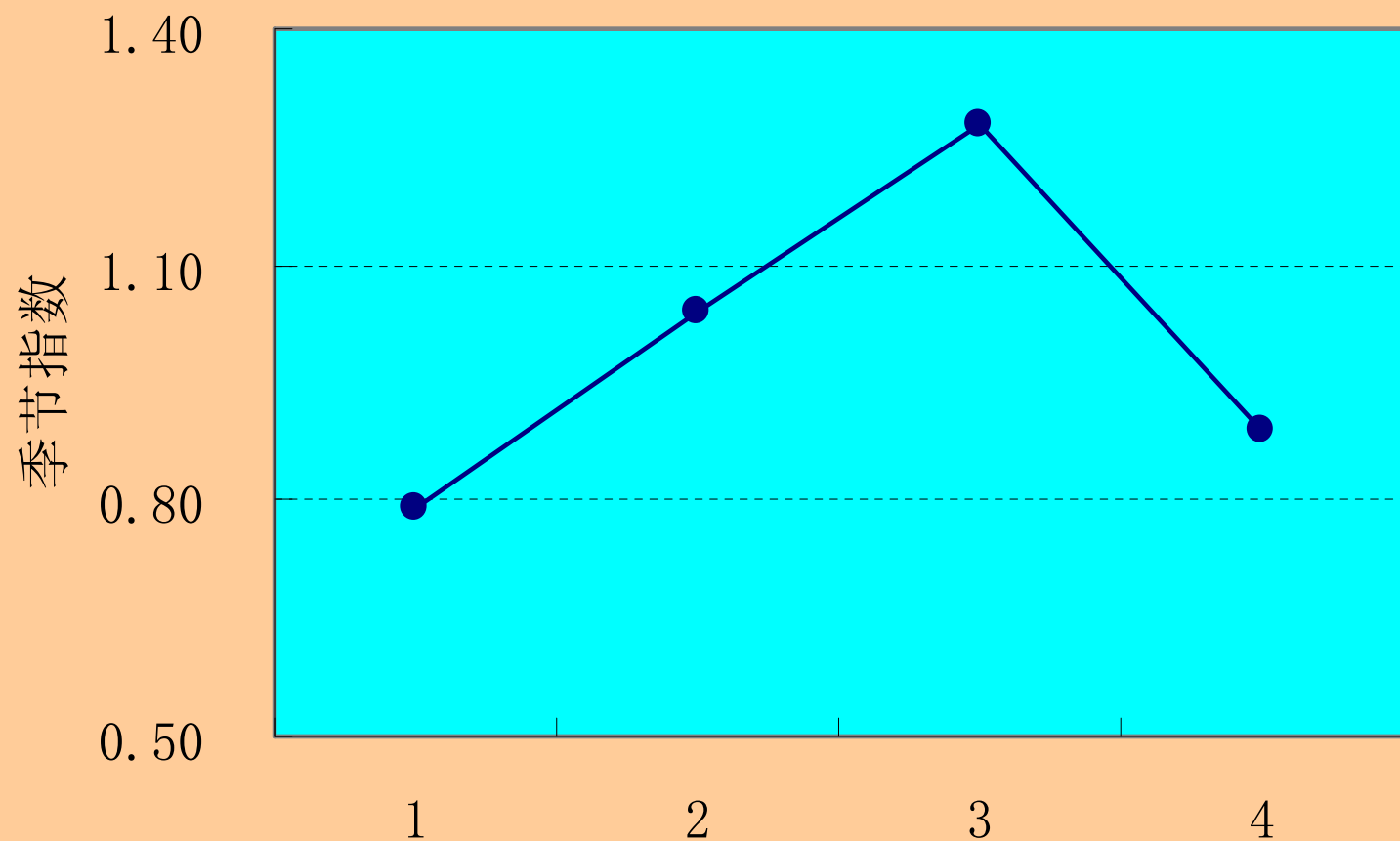
季节指数 (例题分析)

	A	B	C	D	E
	年/季度	时间标号 t	销售量 (Y)	中心化移动平均值 (CMA)	比值 (Y/CMA)
1					
2	2000/1	1	25	—	—
3	2	2	32	—	—
4	3	3	37	30.625	1.2082
5	4	4	26	32.000	0.8125
6	2001/1	5	30	33.375	0.8989
7	2	6	38	34.500	1.1014
8	3	7	42	34.875	1.2043
9	4	8	30	34.875	0.8602
10	2002/1	9	29	36.000	0.8056
11	2	10	39	37.625	1.0365
12	3	11	50	38.375	1.3029
13	4	12	35	38.500	0.9091
14	2003/1	13	30	38.625	0.7767
15	2	14	39	39.000	1.0000
16	3	15	51	39.125	1.3035
17	4	16	37	39.375	0.9397
18	2004/1	17	29	40.250	0.7205
19	2	18	42	40.875	1.0275
20	3	19	55	41.250	1.3333
21	4	20	38	41.625	0.9129
22	2005/1	21	31	41.625	0.7447
23	2	22	43	41.875	1.0269
24	3	23	54	—	—
25	4	24	41	—	—

季节指数 (例题分析)

	A	B	C	D	E
1	年份	季 度			
2		1	2	3	4
3	2000	—	—	1.2082	0.8125
4	2001	0.8989	1.1014	1.2043	0.8602
5	2002	0.8056	1.0365	1.3029	0.9091
6	2003	0.7767	1.0000	1.3035	0.9397
7	2004	0.7205	1.0275	1.3333	0.9129
8	2005	0.7447	1.0269	—	—
9	合计	3.9464	5.1924	6.3522	4.4344
10	平均	0.7893	1.0385	1.2704	0.8869
11	季节指数	0.7922	1.0424	1.2752	0.8902

季节指数 (例题分析)



啤酒销售量的季节变动

季度

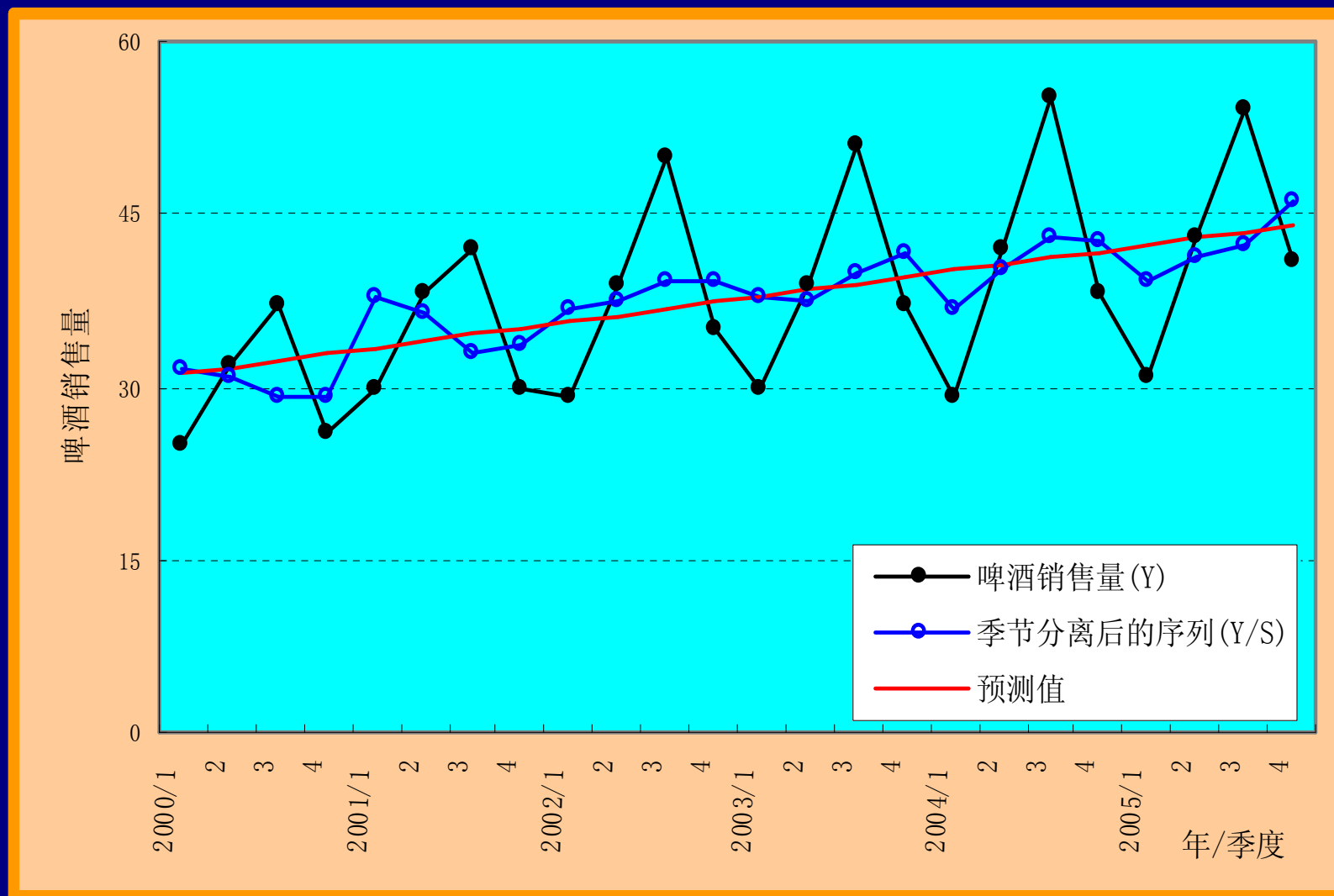
分离季节因素

1. 将原时间序列除以相应的季节指数

$$\frac{Y}{S} = \frac{T \times S \times I}{S} = T \times I$$

2. 季节因素分离后的序列反映了在没有季节因素影响的情况下时间序列的变化形态

季节性及其分离图



统计学
STATISTICS
(第三版)

建立预测模型并进行预测

作者：张占贞，青岛科技大学经济与管理学院

线性趋势模型及预测

1. 根据分离季节性因素的序列确定线性趋势方程

$$\hat{Y}_t = 30.6067 + 0.5592t$$

2. 根据趋势方程进行预测

- 该预测值不含季节性因素，即在沒有季节因素影响情况下的预测值

3. 计算最终的预测值

- 将回归预测值乘以相应的季节指数

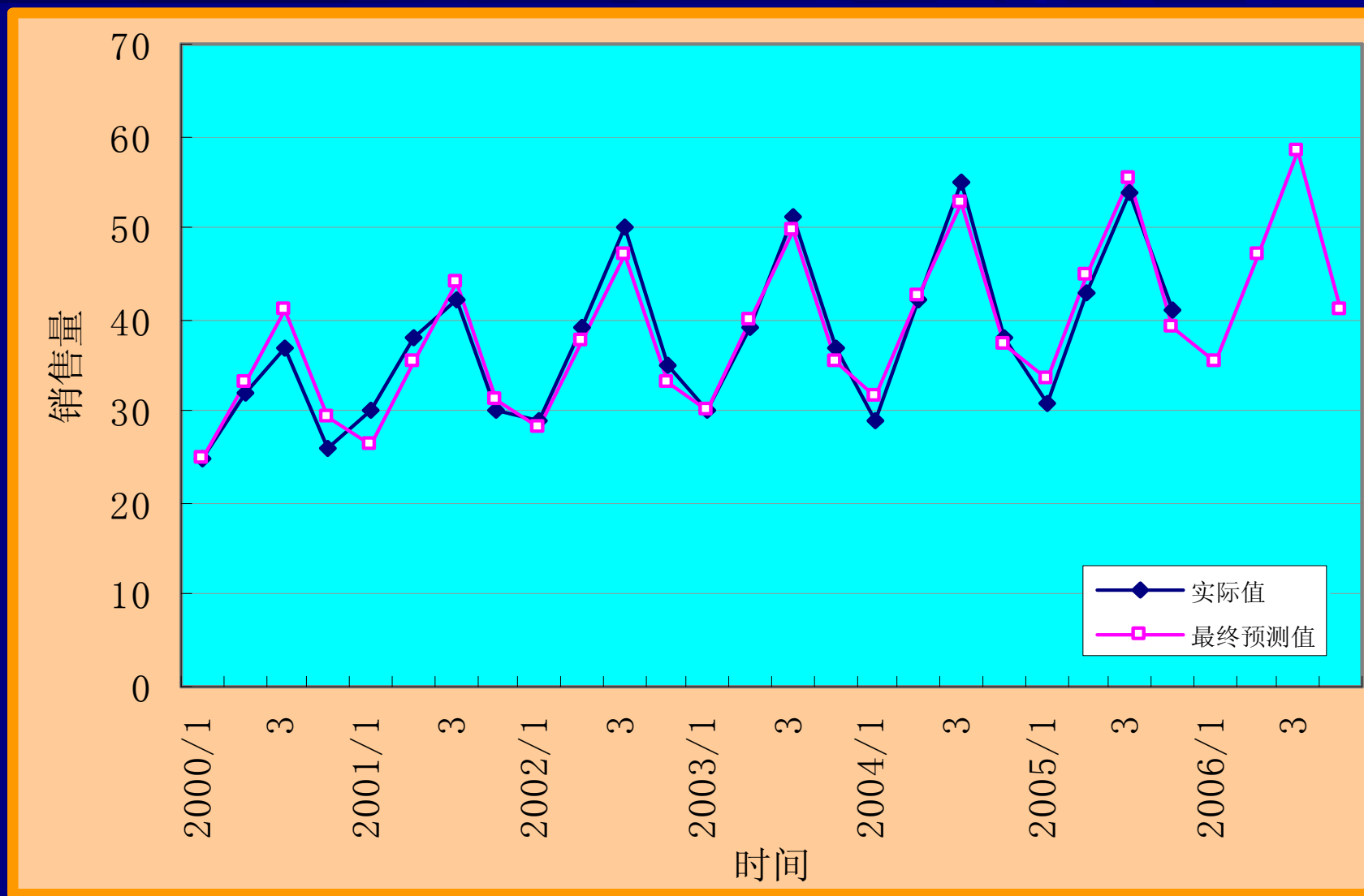
线性趋势预测和最终预测值 (例题分析)

	A	B	C	D	E	F	G	H
1	年/季	时间 编号	啤酒销 售量 (Y)	季节指 数 (S)	季节分离后 的序列 (Y/S)	回归预测值	最终预测值	预测误差
2	(1)	(2)	(3)	(4)	(5) = (3) / (4)	(6)	(7) = (6) × (4)	(8) = (3) - (7)
3	2000/1	1	25	0.7922	31.56	31.17	24.69	0.31
4	2	2	32	1.0424	30.70	31.73	33.07	-1.07
5	3	3	37	1.2752	29.01	32.28	41.17	-4.17
6	4	4	26	0.8902	29.21	32.84	29.24	-3.24
7	2001/1	5	30	0.7922	37.87	33.40	26.46	3.54
8	2	6	38	1.0424	36.46	33.96	35.40	2.60
9	3	7	42	1.2752	32.94	34.52	44.02	-2.02
10	4	8	30	0.8902	33.70	35.08	31.23	-1.23
11	2002/1	9	29	0.7922	36.61	35.64	28.23	0.77
12	2	10	39	1.0424	37.41	36.20	37.73	1.27
13	3	11	50	1.2752	39.21	36.76	46.87	3.13
14	4	12	35	0.8902	39.32	37.32	33.22	1.78
15	2003/1	13	30	0.7922	37.87	37.88	30.01	-0.01
16	2	14	39	1.0424	37.41	38.44	40.06	-1.06
17	3	15	51	1.2752	39.99	38.99	49.73	1.27
18	4	16	37	0.8902	41.56	39.55	35.21	1.79
19	2004/1	17	29	0.7922	36.61	40.11	31.78	-2.78
20	2	18	42	1.0424	40.29	40.67	42.40	-0.40
21	3	19	55	1.2752	43.13	41.23	52.58	2.42
22	4	20	38	0.8902	42.69	41.79	37.20	0.80
23	2005/1	21	31	0.7922	39.13	42.35	33.55	-2.55
24	2	22	43	1.0424	41.25	42.91	44.73	-1.73
25	3	23	54	1.2752	42.35	43.47	55.43	-1.43
26	4	24	41	0.8902	46.06	44.03	39.19	1.81

2006年预测值 (例题分析)

	A	B	C	D	E
1	年/季	时间编号	季节指数(S)	回归预测值	最终预测值
2	2006/1	25	0.79	44.59	35.32
3	2	26	1.04	45.15	47.06
4	3	27	1.28	45.71	58.28
5	4	28	0.89	46.26	41.18

实际值和最终预测值图



本章小节

1. 时间序列的分解
2. 时间序列的描述性分析
3. 时间序列的预测程序
4. 平稳序列的预测
5. 有趋势序列的分析和预测
6. 复合型序列的分解预测