

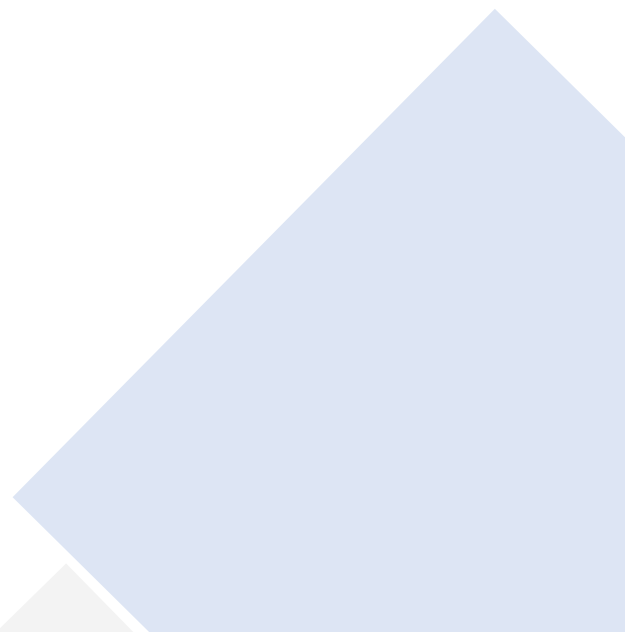
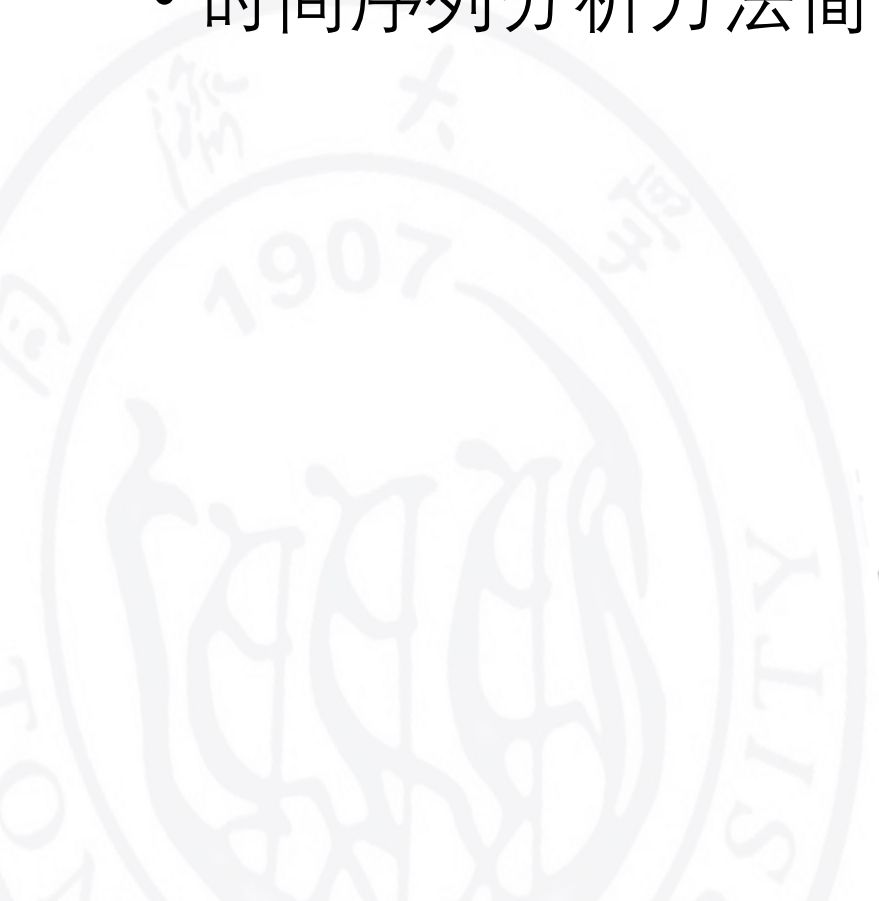
# 统计分析与建模

高珍

[gaozhen@tongji.edu.cn](mailto:gaozhen@tongji.edu.cn)

# 时间序列分析

- 时间序列定义
- 时间序列分析方法简介



# 最早的时间序列分析

- ❖ 7000年前，古埃及人把尼罗河涨落的情况逐天记录下来，就构成所谓的时间序列。
- ❖ 对这个时间序列长期的观察使他们发现尼罗河的涨落非常有规律。当天狼星第一次和太阳同时升起的那一天之后，再过两百天左右，尼罗河就开始泛滥，泛滥期将持续七、八十天，洪水过后，土地肥沃，随意播种就会有丰厚的收成。



- ❖ 由于掌握了尼罗河泛滥的规律，使得古埃及的农业迅速发展，解放出大批的劳动力去从事非农业生产，从而创建了埃及灿烂的史前文明。

# 基本概念

- 时间序列

- 按照时间的顺序把一个随机事件变化发展的过程记录下来就构成了一个时间序列

- ❖ 随机时间序列

按时间顺序排列的一组随机变量

$$X_1, X_2, \dots, X_t, \dots$$

表示一个随机事件的时间序列，简记为  $\{X_t, t \in T\}$ 。

- ❖ 观察值序列

用  $x_1, x_2, \dots, x_n$  或者  $\{x_t, t=1, 2, \dots, n\}$  表示该随机序列的  $n$  个有序观察值，称之为序列长度为  $n$  的观察值序列。

# 基本概念

- 时间序列分析

- 对时间序列进行观察、研究，找寻它变化发展的规律，预测它将来的走势就是时间序列分析

- ❖ 观察值序列是随机时间序列的一个实现（一次抽样过程）

- ❖ 研究目的：揭示随机时序的性质（揭示总体分布、边缘分布、联合分布、均值、方差等）

- ❖ 研究方法：通过观察值序列的性质进行推断（通过可观测样本推断总体参数或者分布）



# 案例

❖ 目标：研究北京市城镇居民消费支出的发展变化规律

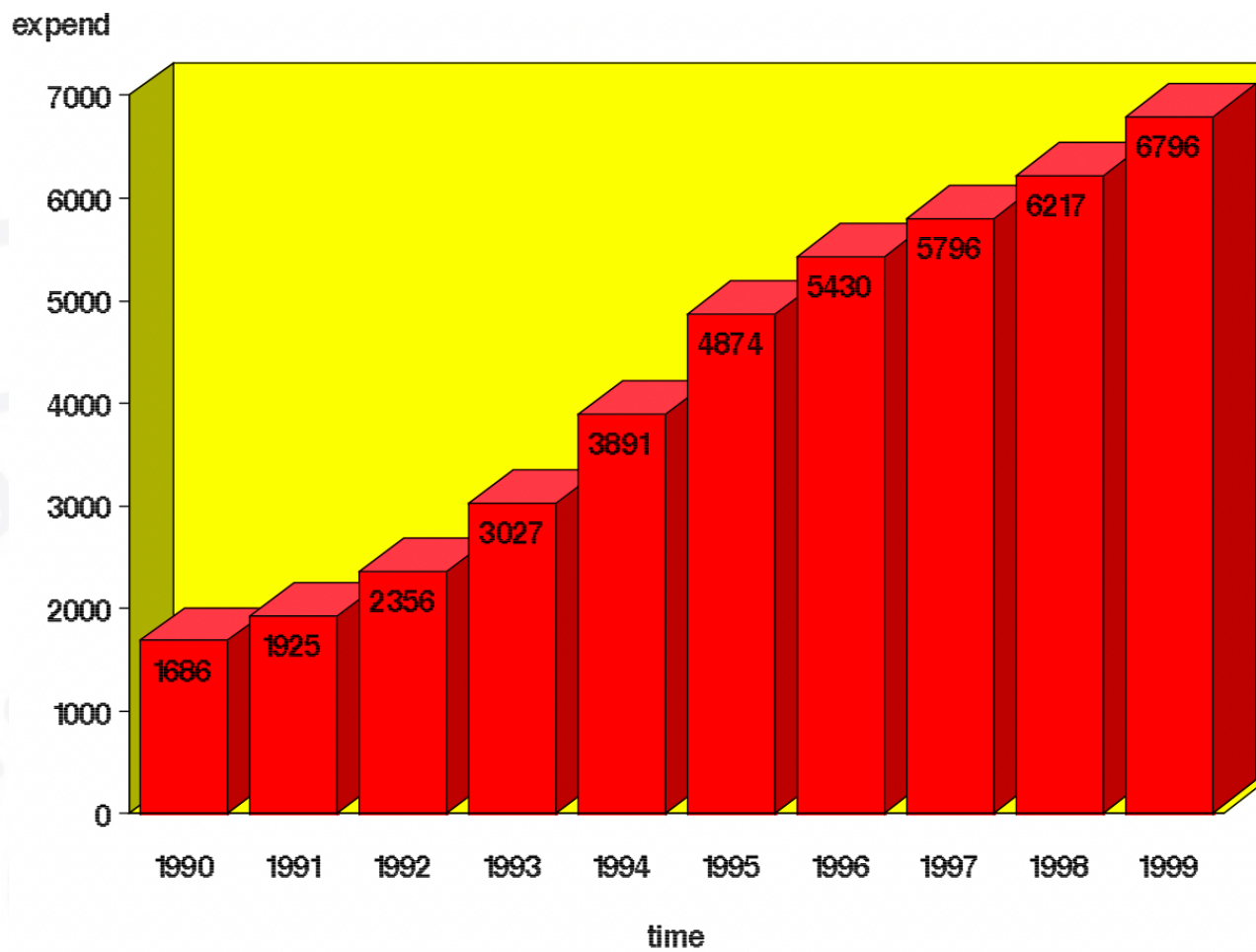
❖ 那么每一年北京市城镇居民的消费支出就构成了一个随机时间序列

$$\cdots, X_{1990}, X_{1991}, \cdots, X_{2017}, X_{2018}, \cdots$$

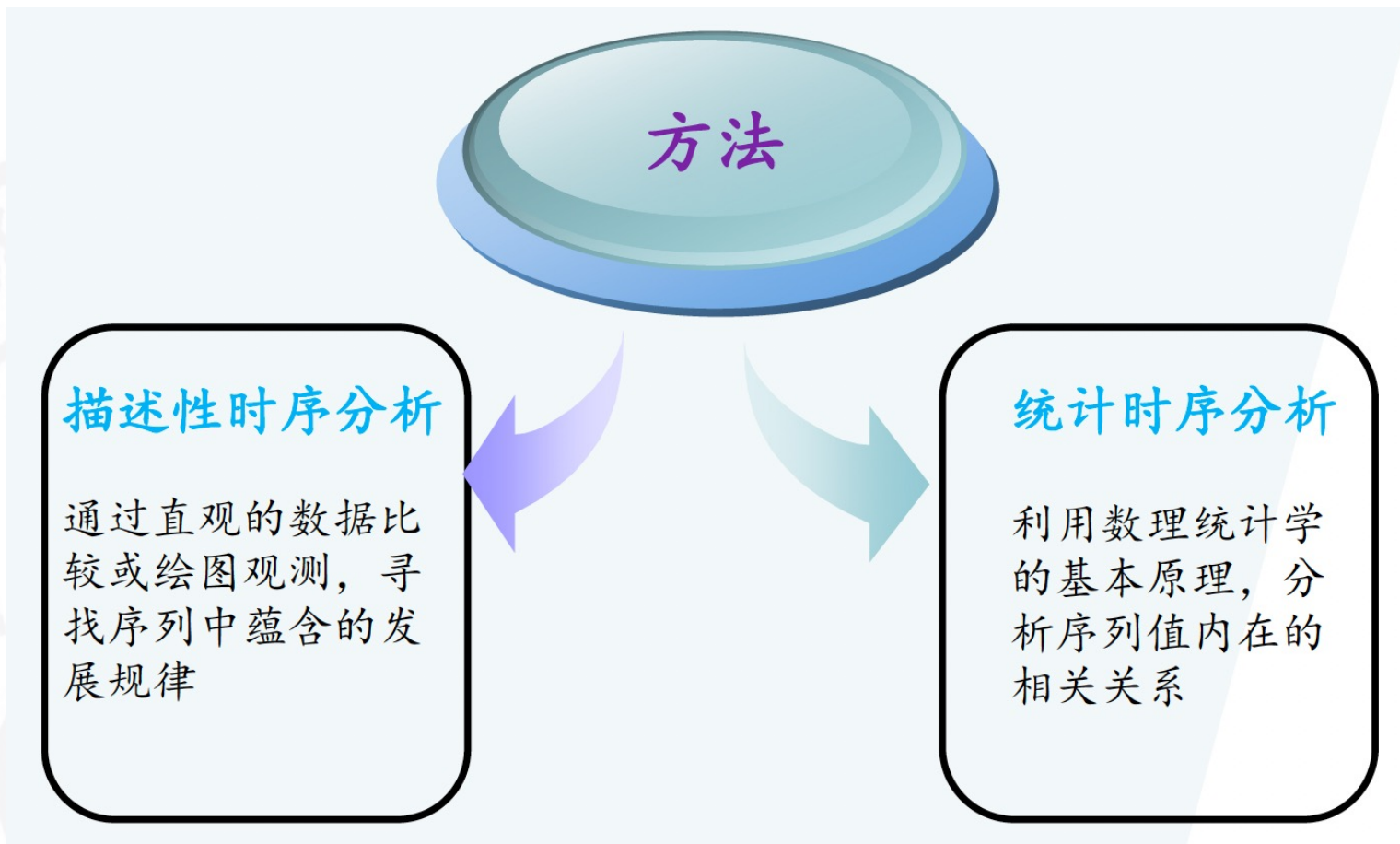
❖ 所获得的某些年份的北京市城镇居民的消费支出序列就是一个观察值序列。例如：北京市1990-1999年每年消费支出（单位：亿元）

1686, 1925, 2356, 3027, 3891, 4874, 5430, 5796,  
6217, 6796

# 案例



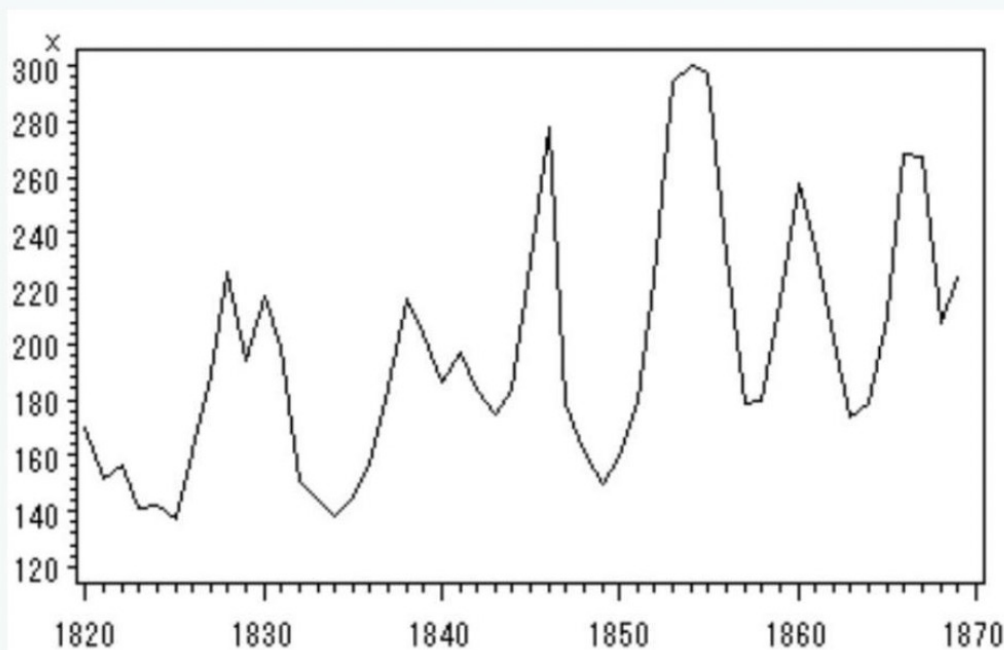
# 分析方法





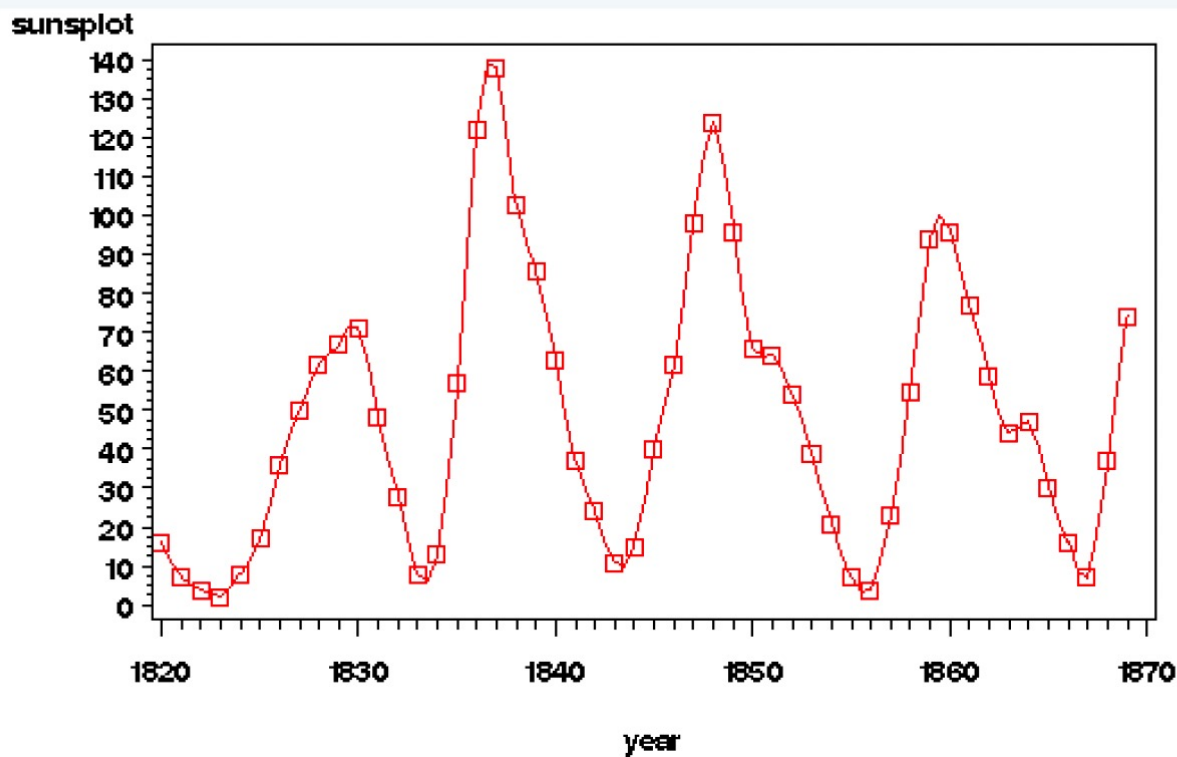
# 更多案例

- ❖ 贝弗里奇(Beveridge)小麦价格指数序列，它由1500-1869年逐年估计的小麦价格构成，可以清晰地看到该序列有一个13年左右的周期。



# 更多案例


- ❖ 德国业余天文学家施瓦贝 (S.H.Schwabe) 发现太阳黑子的活动具有11年左右的周期



# 统计时序分析方法的产生

❖ **描述性时序分析法的局限性**：它依赖于人为主观性，只能展示非常明显的规律性。

❖ **统计时序分析方法的产生**

靠直接观察难以捕捉到规律  用数理统计学原理来分析数据，寻找规律（统计时序分析方法）

# 主要方法解读

## 1, 频域分析

- 频谱分析或谱分析
- 假设任何一种无趋势的时间序列都可以分解为若干不同频率的周期波动
- 傅立叶变换、最大熵谱估计理论
- 方法复杂、结果抽象

## 2, 时域分析

# 时域分析方法

- ❖ 时域分析方法主要是从**序列自相关**的角度揭示时间序列的发展规律。
- ❖ **基本思想**：事件的发展通常都具有一定的惯性，这种惯性用统计的语言来描述就是序列值之间存在着一定的相关关系，且这种相关关系通常具有某种统计规律。
- ❖ **目标**：寻找出序列值之间相关关系的统计规律，并拟合出适当的数学模型来描述这种规律，进而利用这个拟合模型预测序列未来的走势
- ❖ **优点**：理论基础扎实，操作步骤规范，分析结果易于解释。





谢谢!

高珍

[gaozhen@tongji.edu.cn](mailto:gaozhen@tongji.edu.cn)

