

法律声明

□ 本课件包括：演示文稿，示例，代码，题库，视频和声音等，小象学院拥有完全知识产权的权利；只限于善意学习者在本课程使用，不得在课程范围外向任何第三方散播。任何其他人或机构不得盗版、复制、仿造其中的创意，我们将保留一切通过法律手段追究违反者的权利。

□ 课程详情请咨询

■ 微信公众号：小象

■ 新浪微博：ChinaHadoop



目 录

- 第一章 时间序列分析简介
- 第二章 时间序列的预处理
- 第三章 平稳时间序列分析
- 第四章 非平稳序列的确定性分析
- 第五章 非平稳序列的随机分析
- 第六章 多元时间序列分析

第一章 时间序列分析简介

主讲教师 周仕君

本章结构

- 引言
- 时间序列的定义
- 时间序列分析方法简介
- 时间序列分析的软件

第一章 时间序列分析简介

1.1 引言

□ 最早的时间序列分析可以追溯到7000年前的古埃及。

■ 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。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，从而创建了埃及灿烂的史前文明。

尼罗河泛滥期



□ 按照时间的顺序把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。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。

1.2 时间序列的定义

□ 随机序列:按时间顺序排列的一组随机变量

$$\cdots, X_1, X_2, \cdots, X_t, \cdots$$

□ 观察值序列: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 x_1, x_2, \cdots, x_t

□ 随机序列和观察值序列的关系

■ 观察值序列是随机序列的一个实现

■ 我们研究的目的是想揭示随机时序的性质

■ 实现的手段都是通过观察值序列的性质进行推断

案例

- 我们想研究全国高校招生人数的发展变化规律
- 那么每一年全国普通高等学校的招生人数就构成了一个随机序列

$$\cdots, X_{1990}, X_{1991}, \cdots, X_{2012}, X_{2013}, \cdots$$

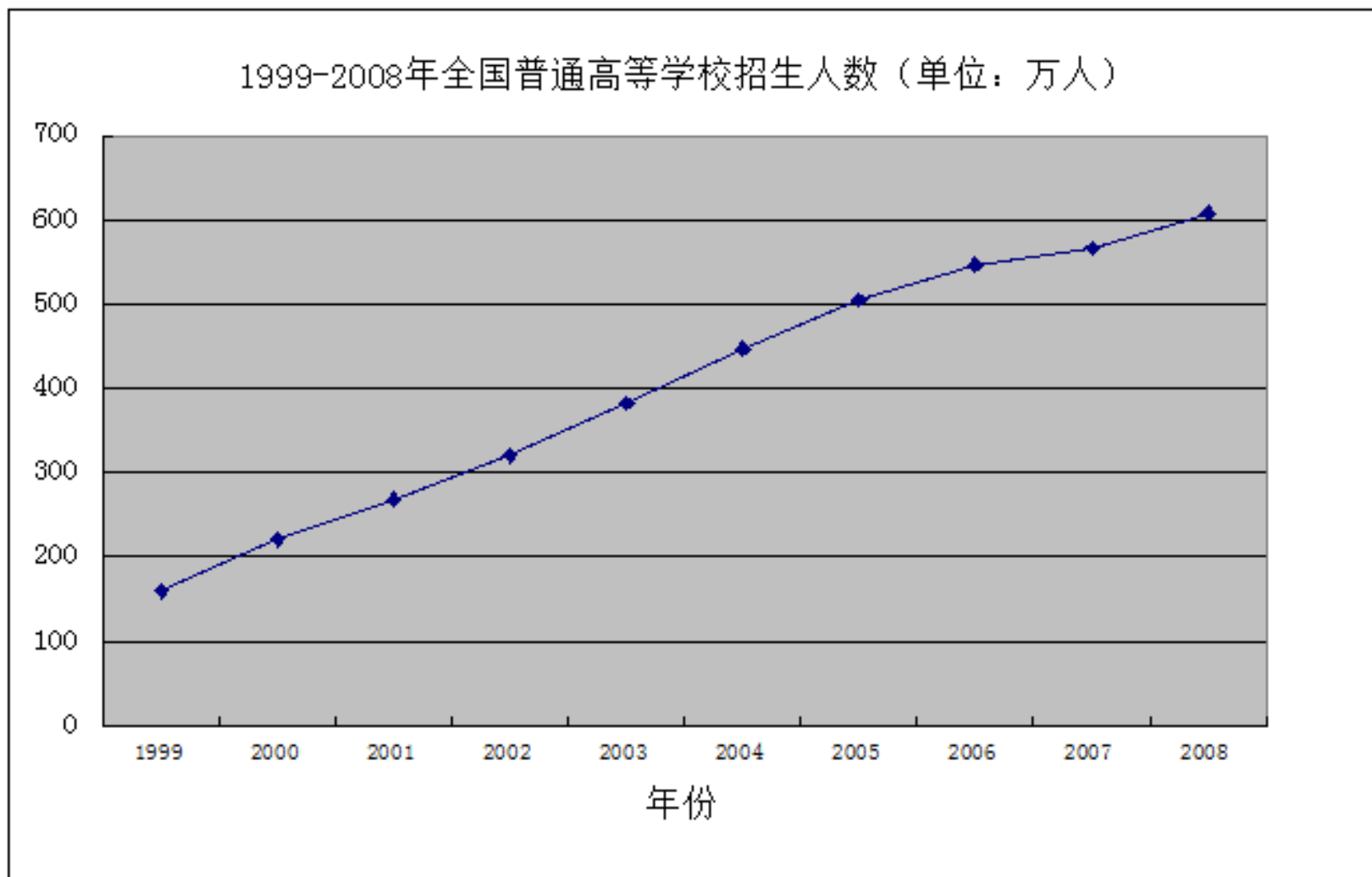
- 通过统计，得到的1999—2008年全国高等学校的招生人数序列就构成一个序列长度为10的观察值序列

我国1999~2008年全国普通高等学校招生人数（单位：万人）

159.7, 220.6, 268.3, 320.5, 382.2

447.3, 504.5, 546.1, 565.9, 607.7

1999-2008年全国普通高等学校招生人数（单位：万人）



1.3 时间序列分析方法

□ 描述性时序分析

□ 统计时序分析

描述性时序分析

- 通过直观的数据比较或绘图观测，寻找序列中蕴含的发展规律，这种分析方法就称为描述性时序分析
- 描述性时序分析方法具有操作简单、直观有效的特点，它通常是人们进行统计时序分析的第一步。

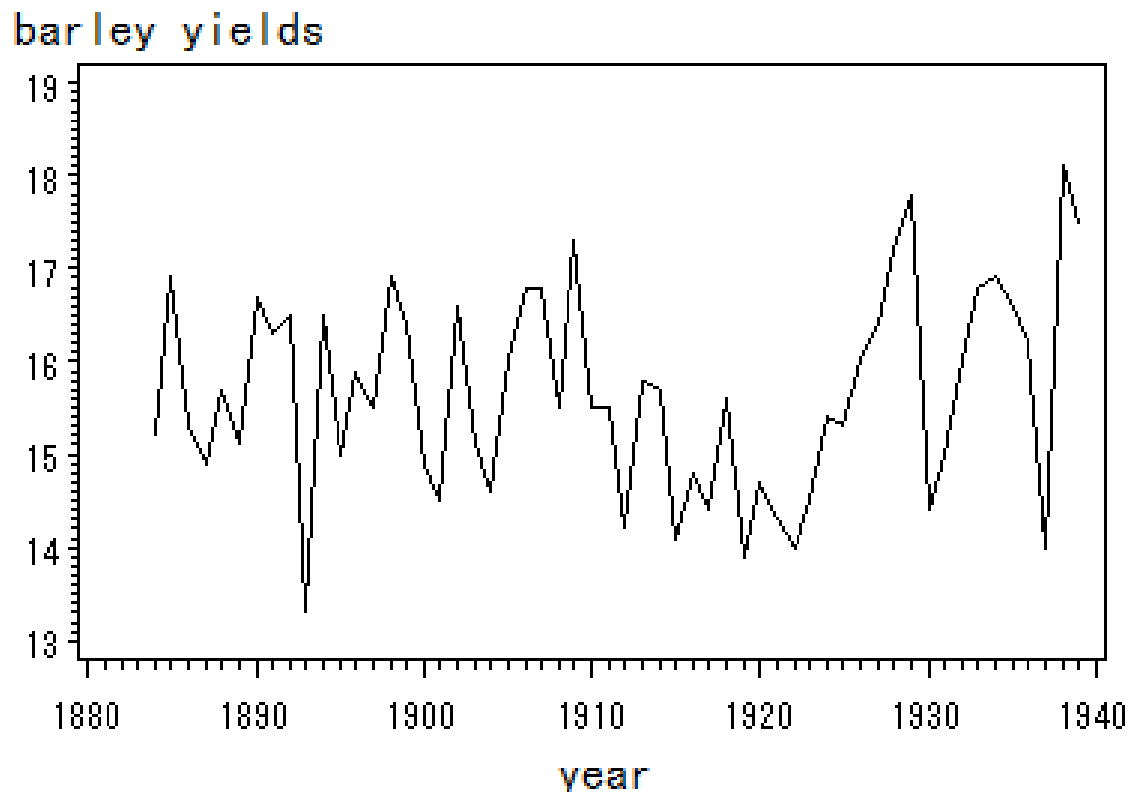
描述性时序分析案例

- 《史记·货殖列传》记载，早在我国春秋战国时期，范蠡和计然就提出我国农业生产具有“六岁穰，六岁旱，十二岁一大饥”的自然规律。
- 《越绝书·计倪内经》则描述得更加详细：“太阴三岁处金则穰，三岁处水则毁，三岁处木则康，三岁处火则旱.....天下六岁一穰，六岁一康，凡十二岁一饥。”
- 翻译成现代文就是：“木星绕天空运行，运行三年，如果处于金位，则该年为大丰收年；如果处于水位，则该年为大灾年；再运行三年，如果至木位，则该年为小丰收年；如果处于火位，则该年为小灾年，所以天下平均六年一大丰收，六年一小丰收，十二年一大饥荒。”
- 这是2500多年前，我国对农业生产具有3年一小波动，12年左右一个大周期的记录，是一个典型的描述性时间序列分析。

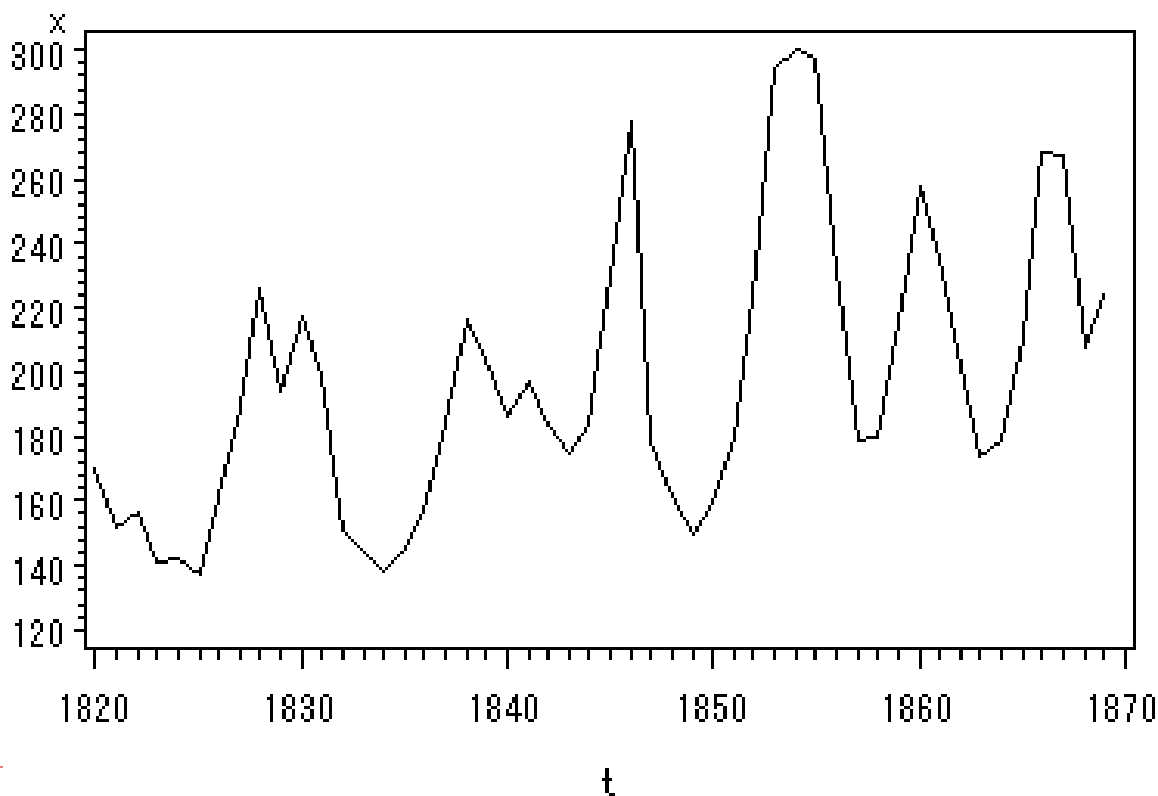
-
- 范蠡根据“六岁穰，六岁旱，十二岁一大饥”的自然规律提出了我国最早稳定粮价的方法：“平粜法”
 - “夫粜，二十病农，九十病末。末病则财不出，农病则草不辟矣。上不过八十，下不减三十，则农末俱利，平粜齐物，关市不乏，治国之道也。”
 - 这段话的意思是：如果是丰收年，粮食贱卖，会伤害农民种粮的积极性。如果是大灾年，粮食天价，会伤到老百姓的生存。所以要实行“平粜法”。政府应该在粮食丰收时高于最低价购买粮食进行储备，以保护农民的利益；在粮食短缺时，将储备粮食投放市场，以稳定粮价，确保百姓的生存。这是对农民和百姓都有利的政策，是一个国家的治国之道。

- 在范蠡之后2000年，欧洲经济学家在研究欧洲各地粮食产量时发现了类似规律。

1884-1939年苏格兰与威尔士每英亩大麦产量时序图

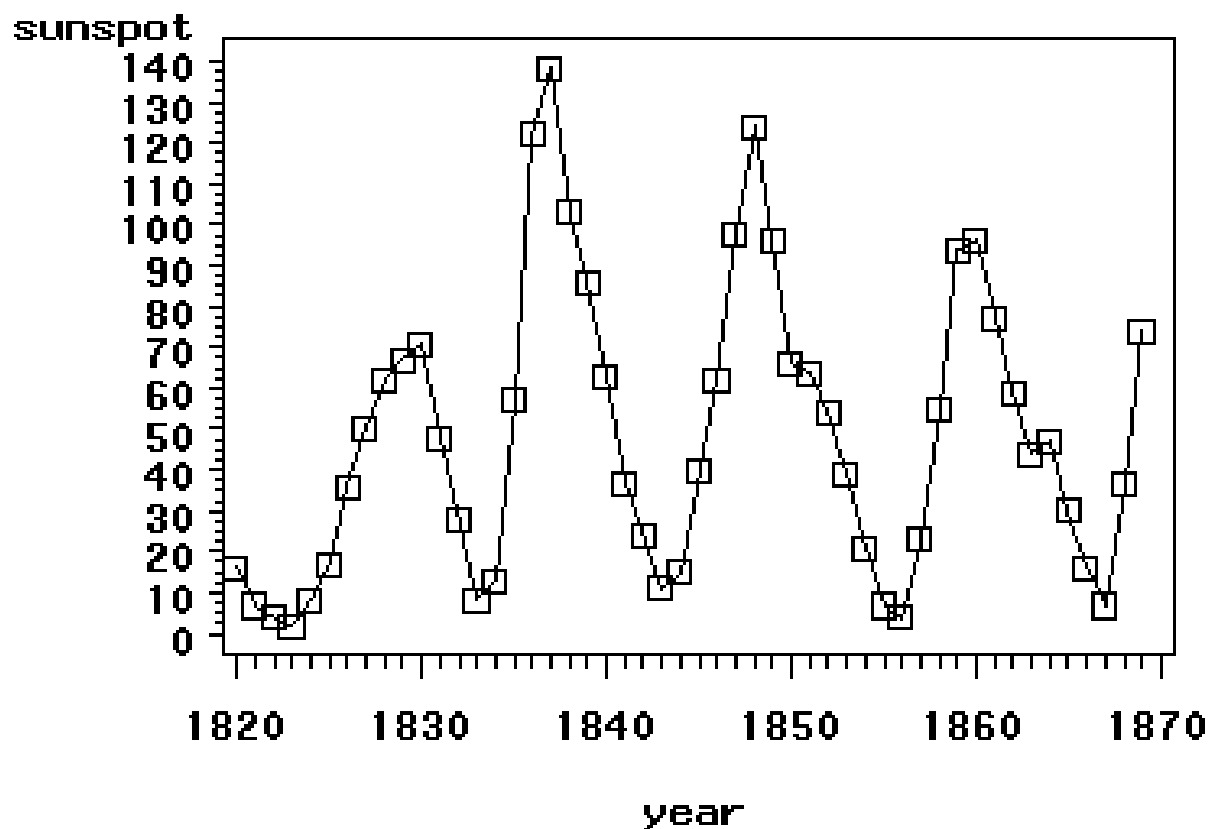


- ❑ 贝弗里奇（Beveridge）小麦价格指数序列，它由1500-1869年逐年估计的小麦价格构成，可以清晰地看到该序列有一个13年左右的周期



描述性时序分析案例

- 德国业余天文学家施瓦尔发现太阳黑子的活动具有11年左右的周期



-
- 太阳黑子的运动周期和农业生产的周期长度非常接近，这引起了英国天文学家、天王星的发现者威廉 赫歇尔（F. W. Herschel）的关注。最后他发现当太阳黑子变少时，地球上的雨量也会减少。所以在没有良好人工灌溉技术的时代，农业生产会呈现出和太阳黑子近似的变化周期。
 - 我们没有采用任何复杂的模型或分析方法，仅仅是沿着时间顺序收集数据，描述和呈现序列的波动，就了解到小麦产量的周期波动特征，以及产生该周期波动的气候成因及该周期对价格的影响。

-
- 通过直观的数据比较或绘图观测，寻找序列中蕴含的发展规律，这种分析方法就称为描述性时序分析
 - 描述性时序分析方法具有操作简单、直观有效的特点，它通常是人们进行统计时序分析的第一步。
 - 局限性：它只能展示非常明显的规律性。
 - 而在金融、保险、法律、人口、心理学等社会科学研究领域，随机变量的发展通常会呈现出非常强的随机性，想通过对序列简单的观察和描述，总结出随机变量发展变化的规律，并准确预测出它们将来的走势通常是非常困难的。

统计时序分析

□ 频域分析方法

□ 时域分析方法

频域分析方法

□ 原理

- 假设任何一种无趋势的时间序列都可以分解成若干不同频率的周期波动

□ 发展过程

- 早期的频域分析方法借助富里埃分析从频率的角度揭示时间序列的规律
- 后来借助了傅里叶变换，用正弦、余弦项之和来逼近某个函数
- 20世纪60年代，引入最大熵谱估计理论，进入现代谱分析阶段

□ 特点

- 非常有用的动态数据分析方法，但是由于分析方法复杂，结果抽象，有一定的使用局限性

时域分析方法

□ 原理

- 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，这种相关关系通常具有某种统计规律。

□ 目的

- 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

□ 特点

- 理论基础扎实，操作步骤规范，分析结果易于解释，是时间序列分析的主流方法

时域分析方法的分析步骤

- 考察观察值序列的特征
- 根据序列的特征选择适当的拟合模型
- 根据序列的观察数据确定模型的口径
- 检验模型，优化模型
- 利用拟合好的模型来推断序列其它的统计性质或预测序列将来的发展

时域分析方法的发展过程

- 基础阶段
- 核心阶段
- 完善阶段

基础阶段

□ G.U.Yule

- 英国数学家。1927年，Yule提出用线性回归方程来模拟一个时间序列，这是最早的AR模型。

□ G.T.Walker

英国数学家，天文学家。

1931年，Walker利用Yule的分析方法 研究了衰减正弦时间序列，得出Yule-Walker方程



核心阶段

□ G.E.P.Box和 G.M.Jenkins

- 1970年，他们出版了《Time Series Analysis Forecasting and Control》一书
- 书中，他们系统地阐述了ARIMA模型的识别、估计、检验及预测的原理及方法。这些知识现在被称为经典时间序列分析方法，是时域分析方法的核心内容。
- 为了纪念Box和Jenkins对时间序列发展的特殊贡献，现在人们也常把ARIMA模型称为Box-Jenkins模型。

□ ARIMA模型的实质

- 单变量、同方差场合的线性模型

完善阶段

□ 异方差场合

- Robert F.Engle, 1982年, ARCH模型
- Bollerslov, 1985年, GARCH模型
- Nelson等人提出了GARCH模型的多种衍生模型

□ 多变量场合

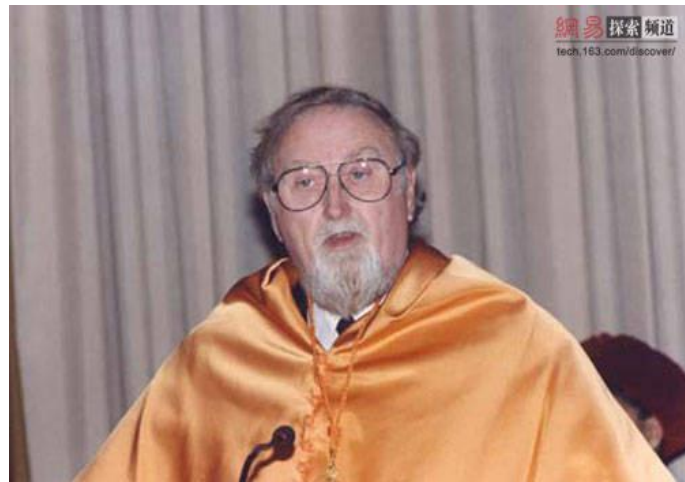
- C.Granger, 1987年提出了协整（co-integration）理论

□ 非线性场合

- 汤家豪等, 1980年, 门限自回归模型
- C.Granger, 1978年, 双线性模型



Robert F.Engle



C.Granger

时间序列分析的软件

□ R语言

疑问

□ 问题答疑：<http://www.xxwenda.com/>

■ 可邀请老师或者其他人回答问题

联系我们

小象学院：互联网新技术在线教育领航者

- 微信公众号：小象
- 新浪微博：ChinaHadoop

