

时间序列分析(初级)

时间序列分析简介

陈垚翰

安徽大学 大数据与统计学院



安徽大学
Anhui University



本章结构

- 引言
- 时间序列的定义
- 时间序列分析的方法和工具

引言

引言

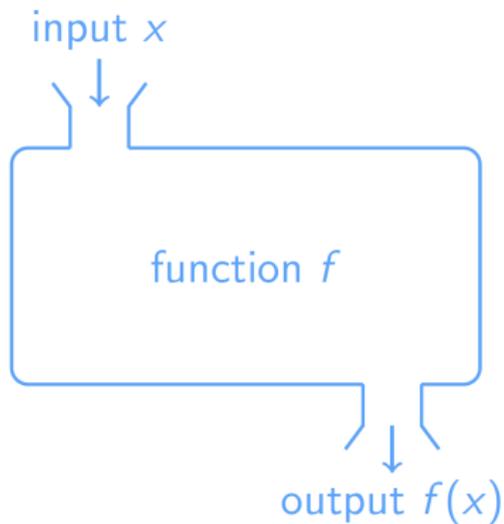
- 最早的时间序列分析可以追溯到 7000 年前的古埃及。

- 最早的时间序列分析可以追溯到 7000 年前的古埃及。
- 《史记·货殖列传》：“六岁穰，六岁旱，十二岁一大饥”。
- 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。

时间序列的定义

函数

函数 是将一个集合的元素与另一个集合的元素建立关系的规则。函数有输入有输出。



随机变量

- 随机变量是函数, 将随机结果映射为数值。
- 当投掷一次硬币, 以 ‘1’ 表示正面, ‘0’ 表示背面。以这种方式, 我们有 随机变量 $X(\omega) \mapsto \{0, 1\}$, 其中 ω 属于 样本空间 $\Omega = \{\text{正面}, \text{背面}\}$ 。
- Ω 是包括 所有随机结果 的集合。
- 随机变量 $X(\omega)$ 是一个 取实数值 定义在 Ω 上的函数。

- Ω 的子集构成的集合 \mathcal{F} 。
- 如何使 \mathcal{F} 完备? \mathcal{F} 应满足:
 - 如果 $A \in \mathcal{F}$, 则 $A^c \in \mathcal{F}$;
 - 如果 $A, B \in \mathcal{F}$, 则 $A \cap B \in \mathcal{F}, A \cup B \in \mathcal{F}, A \cup B^c \in \mathcal{F}, B \cup A^c \in \mathcal{F}, A \cap B^c \in \mathcal{F}, B \cap A^c \in \mathcal{F}$ 。
- 在高级课程中 \mathcal{F} 称作 **σ -域 (σ -field)**。

时间序列的定义

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

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

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

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

- 观测值序列: 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

$$x_1, x_2, \cdots, x_n$$

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

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

- 观测值序列: 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

$$x_1, x_2, \cdots, x_n$$

- 随机序列和观测值序列的关系

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

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

- 观测值序列: 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

$$x_1, x_2, \cdots, x_n$$

- 随机序列和观测值序列的关系
 - 观测值序列是随机序列的一个实现

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

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

- 观测值序列: 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

$$x_1, x_2, \cdots, x_n$$

- 随机序列和观测值序列的关系
 - 观测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

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

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

- 观测值序列: 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

$$x_1, x_2, \cdots, x_n$$

- 随机序列和观测值序列的关系
 - 观测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测值序列的性质进行推断

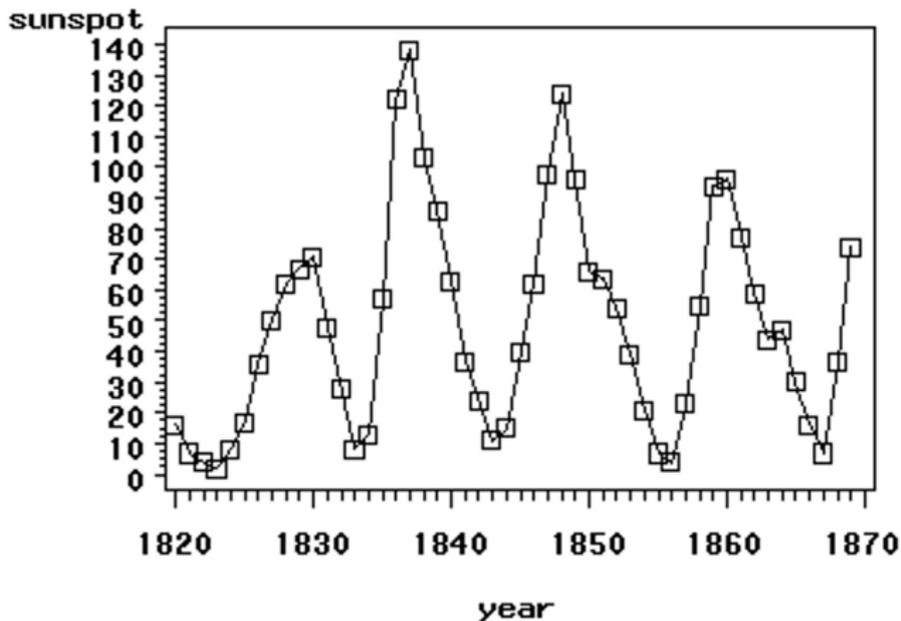
时间序列分析的方法和工具

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

- 太阳黑子活动 11 年左右的周期性



统计时序分析

- 时域分析方法
- 频域分析方法

- 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。

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。

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。

- 考察观察值序列的特征

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

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

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

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

- G.U.Yule
 - 1927 年, AR 模型

基础阶段

- G.U.Yule
 - 1927 年, AR 模型
- G.T.Walker
 - 1931 年, MA 模型, ARMA 模型

- G.E.P.Box 和 G.M.Jenkins
 - 1970 年, 出版《Time Series Analysis Forecasting and Control》
 - 提出 ARIMA 模型 (Box-Jenkins 模型)

- G.E.P.Box 和 G.M.Jenkins

- 1970 年, 出版《Time Series Analysis Forecasting and Control》
- 提出 ARIMA 模型 (Box-Jenkins 模型)
- Box-Jenkins 模型实际上是主要运用于单变量、同方差场合的线性模型

- 异方差场合
 - Robert F. Engle, 1982 年, ARCH 模型
 - Tim Bollerslev, 1986 年, GARCH 模型

- 异方差场合
 - Robert F. Engle, 1982 年, ARCH 模型
 - Tim Bollerslev, 1986 年, GARCH 模型
- 多变量场合
 - C. Granger, 1987 年, 协整(co-integration)理论

完善阶段

- 异方差场合
 - Robert F. Engle, 1982 年, ARCH 模型
 - Tim Bollerslev, 1986 年, GARCH 模型
- 多变量场合
 - C. Granger, 1987 年, 协整 (co-integration) 理论
- 非线性场合

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

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 **Fourier** 分析从频率的角度揭示时间序列的规律

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{i\omega t} dt$$

- 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
- 主要广泛应用于工程学、物理学、天文学和气象科学等领域,近年来在理论计量经济学的研究中逐渐受到重视。

时间序列分析软件

- 常用软件
 - Matlab, R, SAS, Eviews, Gauss
- 本课程主要使用软件: R
- 关于 R 编程
 - Matloff (2011), Crawley (2013)
 - 此外,《R for Data Science》(<https://r4ds.had.co.nz/>)和 Dirk Eddelbuettel 的网站(<https://dirk.eddelbuettel.com/>)也是很好的 R 编程的学习资料。Dirk Eddelbuettel 作为 Linux Debian 和 R 维护的核心成员, Dirk Eddelbuettel 的网站对编程实践具有较高的启发性。

参考文献

参考文献 I

-  CRAWLEY M J, 2013. **The R book**[M]. John Wiley & Sons.
-  MATLOFF N, 2011. **The Art of R Programming**[M].
William Pollock.