

时间序列分析

CH1 & CH2

姓 名:

学 号:

班 级: 信计 1901

江南大学

理学院

2022 年 3 月 9 日

1 CH1 习题

1.1 什么是时间序列？请收集几个生活中的观察值序列。

答：时间序列是指将同一统计指标的数值按其发生的时间先后顺序排列而成的数列。

在 espn 网站上可以找到湖人队对阵鹈鹕队的预测胜率随时间变化的数据，这是时间序列，图像如下：

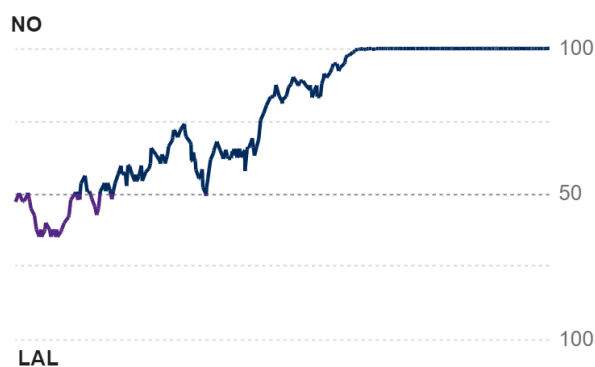


图 1: Win Probability

同样，湖人队与鹈鹕队的比分随时间变化也是时间序列，图像如下：

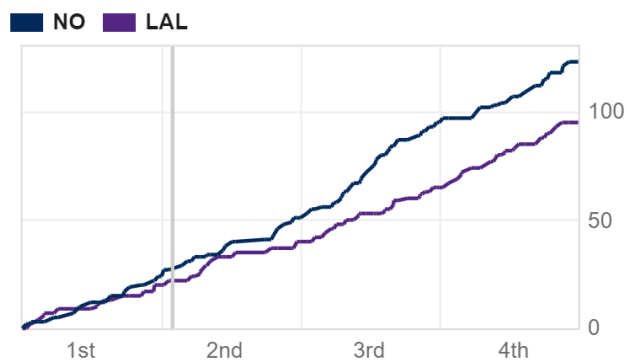


图 2: Game Flow

1.2 时域方法的特点是什么？

答：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。

2 CH2 习题

若序列长度为 100, 前 12 个样本自相关系数如下: $\rho_1 = 0.02, \rho_2 = 0.05, \rho_3 = 0.10, \rho_4 = -0.02, \rho_5 = 0.05, \rho_6 = 0.01, \rho_7 = 0.12, \rho_8 = -0.06, \rho_9 = 0.08, \rho_{10} = -0.05, \rho_{11} = 0.02, \rho_{12} = -0.05$ 。该序列能否视为纯随机序列 ($\alpha = 0.05$) ?

解: 假设

$$H_0: \rho_1 = \rho_2 = \cdots = \rho_{12}$$

$$H_1: \text{至少存在某个 } \rho_k \neq 0, 1 \leq k \leq 12$$

计算统计量 $LB = n(n+2) \sum_{k=1}^m \frac{\hat{\rho}_k^2}{n-k}$, 其中 n 为序列长度 100, m 为延迟期数 12, 计算得到 $LB = 4.98953$, $Q = 4.57$ 。

Input

$$\left(\frac{0.02^2}{99} + \frac{0.05^2}{98} + \frac{0.1^2}{97} + \frac{0.02^2}{96} + \frac{0.05^2}{95} + \frac{0.01^2}{94} + \frac{0.12^2}{93} + \frac{0.06^2}{92} + \frac{0.08^2}{91} + \frac{0.05^2}{90} + \frac{0.02^2}{89} + \frac{0.05^2}{88} \right) \times 100 \times 102$$

Result ↗

4.9895313791489153411953537636507433149129663167185994736597884497241137342534...

More digits

图 3: 计算过程

查表知: $\chi_{0.975}^2(12) = 4.40, \chi_{0.95}^2(12) = 5.23$, 因为 $4.40 < Q < LB < 5.23$, 所以 P 值显著大于显著性水平 0.05。所以不能拒绝原假设, 可以认为该序列为纯随机序列。

3 上机题目

1975-1980 年夏威夷岛莫纳罗亚火山每月释放的 CO_2 数据如下：

3.1 绘制该序列时序图，并判断该序列是否平稳

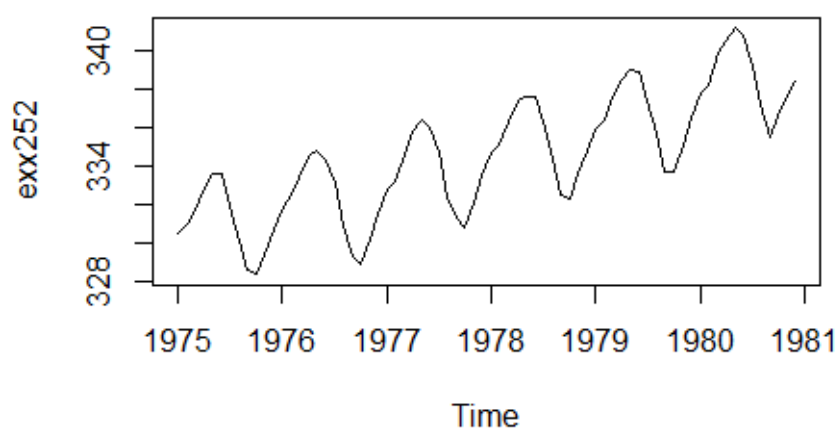


图 4: 时序图

由上图可以看出， CO_2 排量总体逐步上升，且以年为周期呈现出一定的周期性。故该序列是呈现带周期性的单调上升趋势，该序列不平稳。

3.2 计算该序列样本自相关系数 $\hat{\rho}_k$

```
[,1]  
[1,] 1.00000000  
[2,] 0.90750778  
[3,] 0.72171377  
[4,] 0.51251814  
[5,] 0.34982244  
[6,] 0.24689637  
[7,] 0.20309427  
[8,] 0.21020799  
[9,] 0.26428810  
[10,] 0.36433219  
[11,] 0.48471672  
[12,] 0.58456166  
[13,] 0.60197891  
[14,] 0.51841257  
[15,] 0.36856286  
[16,] 0.20671211  
[17,] 0.08138070  
[18,] 0.00135460  
[19,] -0.03247916  
[20,] -0.02709893  
[21,] 0.01123597  
[22,] 0.08274806  
[23,] 0.17010715  
[24,] 0.24319854  
[25,] 0.25252294
```

图 5: 自相关系数

3.3 绘制该样本自相关图，并解释该图形

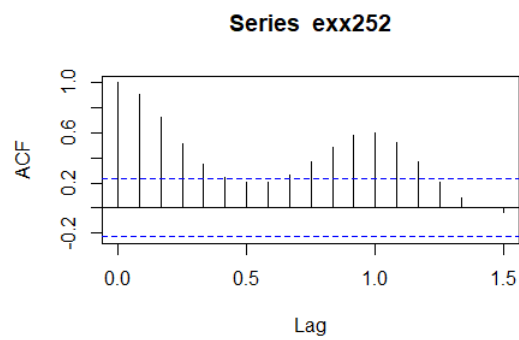


图 6: 自相关图

该图形显示序列自相关系数长期位于零轴的一边, 这是具有单调趋势序列的典型特征, 同时自相关图呈现出明显的正弦波动规律, 这是具有周期变化规律的非平稳序列的典型特征。自相关图显示出带长期递增趋势的周期性性质。

3.4 随机性检验

假设:

$$H_0: \rho_1 = \rho_2 = \cdots = \rho_{24}$$

H_1 : 至少存在某个 $\rho_k \neq 0, 1 \leq k \leq 24$

```
> Box.test(exx252, lag=24, type="Ljung-Box")

Box-Ljung test

data:  exx252
X-squared = 301.25, df = 24, p-value < 2.2e-16
```

图 7: 检验

检验显示, 延迟 24 阶的 LB 统计量的 P 值只有 2.2×10^{-16} , 小于显著水平 0.05, 所以显著拒绝原序列伪随机序列的原假设, 认为该序列为非白噪声序列。

3.5 代码

```
exx252 <- read.table("D:/Study/大三下/时间序列分析/2022_02_23/时间序列分析——基于R (第2版)
习题数据/data/2.2data.txt ", sep="\t")
exx252 <- exx252[,1]
for (i in 2:length(exx252[,1]))
  exx252 <- cbind(exx252, exx252[i,])
exx252 <- as.numeric(exx252)
exx252 <- ts(exx252, start=c(1975,1), frequency = 12)
plot(exx252)
acf(exx252)
acf(exx252, lag.max=24, plot=FALSE)
```