

Time Series Analysis

shake_it

方法

- 描述性时序分析
- 统计时序分析 (频域分析方法 + 时域分析方法)

生成数据

从 2005 年 1 月开始的月度数据。start 指定起始读入时间, frequency 指定序列每年读入的数据频率。

```
price <- c(101, 82, 66, 35, 31, 7)

price <- ts(price, start = c(2005, 1), frequency = 12)
```

例 1-1

读入 1884-1939 年英格兰和威尔士小麦平均亩产量数据 file1.csv

```
x <- read.csv("D:/Documents/UIBE/6/TimeSeries/file1.csv")
head(x)
```

```
##   year yield
## 1 1884  15.2
## 2 1885  16.9
## 3 1886  15.3
## 4 1887  14.9
## 5 1888  15.7
## 6 1889  15.1
```

截取 1925 年之后的数据 subset

```
z <- subset(x, year > 1925, select = yield)
head(z)
```

```
##   yield
## 43  16.0
## 44  16.4
## 45  17.2
## 46  17.8
## 47  14.4
## 48  15.0
```

对 yield 序列进行对数变换, 并将对数序列和原序列值导出, 保存为数据文件 yield.csv。

```
ln_yield <- log(x$yield)
x_new <- data.frame(x, ln_yield) # 新数据框
write.csv(x_new, file = "D:/Documents/UIBE/6/TimeSeries/yield.csv", row.names = F)
```

缺失值插值

R 中缺失值用 NA 表示。常用的插值方法：线性插值和样条插值

```
library(zoo)
a <- 1:7
a[4] <- NA
cat("a: ", a)
```

```
## a:  1 2 3 NA 5 6 7
```

```
y1 <- na.approx(a)
y2<-na.spline(a)
cat(" ", y1, "\n ", y2)
```

```
##    1 2 3 4 5 6 7
```

```
##    1 2 3 4 5 6 7
```