TIME SERIES TP 1

2020-09-21

Contents


```
library(ggplot2)
library(astsa)
library(dplyr)
library(xts)
```

EXERCISE 1

Question 1

Ths $ts(data, start{=}1, end{=}\text{numeric}, frequency{=}1, \dots)$ function is used to create a time series object.

- *data*: a vector or matrix of the observed time series
- *start*: the time of the first observation
- *end*: the time of the last observations , specified in the same way as start.
- $frequency$: the numeber of observations par unit of time

The $diff(\text{data})(x, lag{=}1)$ returns suitably and iterated differences

- x: a numeric vector or a matrix containing the values to be differenced
- *lag*: an integer indicating which lag to use

Question 2

The *choose.file function choose a file interactively from the current repository.

Not very efficient when we use rmarkdown for a final report.

```
if(!exists('euro50'))
{
    euro50<-read.csv('tableEurom.csv',header = TRUE)
}
head(euro50)
```

```
        Date    Open    High     Low   Close Volume Adj.Close
1 2015-09-01 3188.73 3198.86 3188.73 3198.86      0   3198.86
2 2015-08-03 3635.40 3676.75 3218.01 3280.78      0   3280.78
3 2015-07-01 3496.28 3686.58 3294.19 3600.69      0   3600.69
4 2015-06-02 3561.89 3583.82 3424.30 3424.30      0   3424.30
5 2015-05-01 3615.59 3632.94 3615.59 3632.94      0   3632.94
6 2015-04-01 3714.89 3828.78 3615.59 3615.59      0   3615.59
```

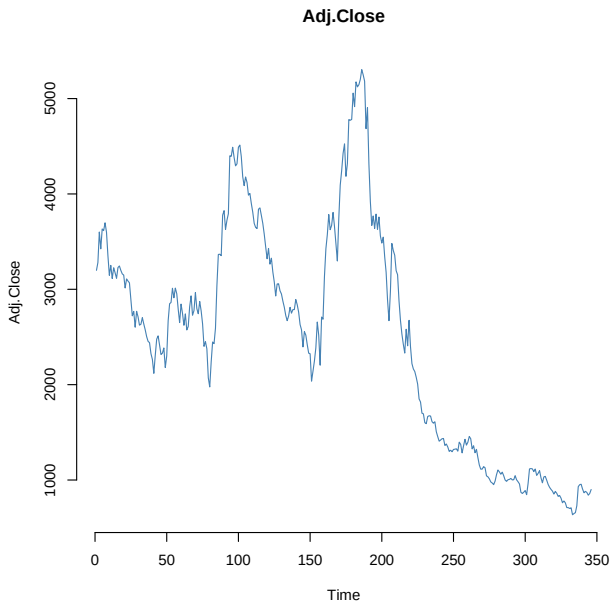
Question 3

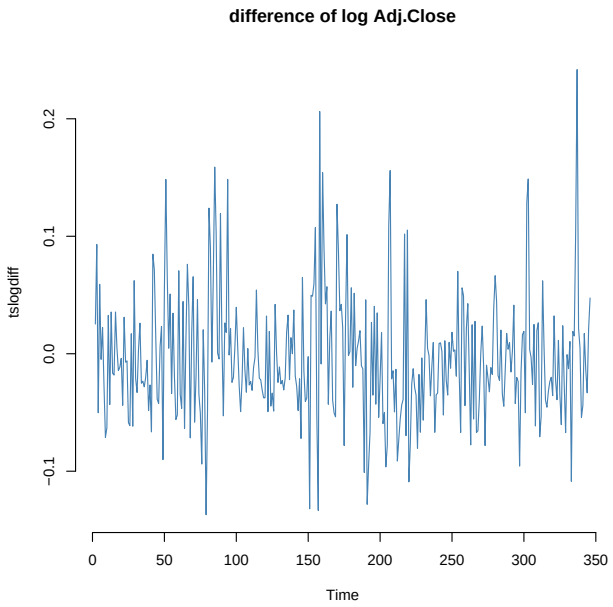
The data contains different informations about a financial index as: the prices at the opening and closing market, the lower and the higher price and so forth.

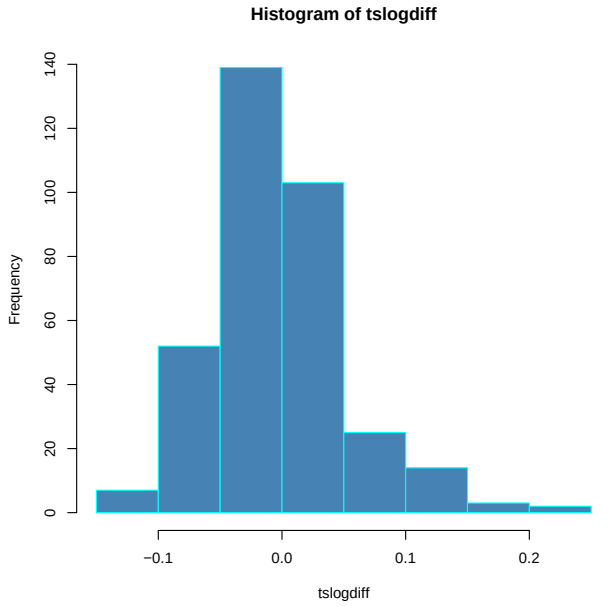
euro50[,7] contains the price at the closing time.

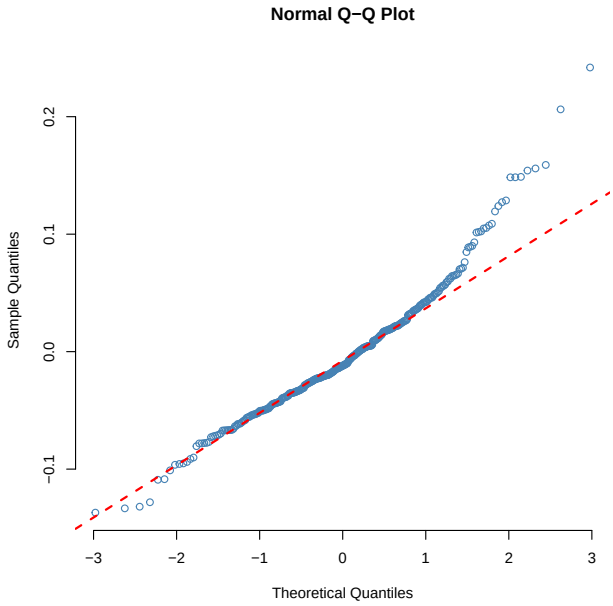
Question 4

```
par(mfrow=c(2,2))
tseries50<-ts(euro50)
tslogdiff<-diff(log(tseries50[,7]))
plot(tseries50[,7],
     main='Adj.Close',
     col='steel blue',
     ylab='Adj.Close',
     frame=FALSE
    )
plot(tslogdiff,
     col='steel blue',
     main='difference of log Adj.Close',
     frame=FALSE
     )
hist(tslogdiff,
     col='steel blue',
     border = 'cyan',
     )
qqnorm(tslogdiff,
       col='steel blue',
       frame=FALSE,
       )
qqline(tslogdiff,
       col='red',
       lty=2,
       lwd=2
       )
```









EXERCISE 2

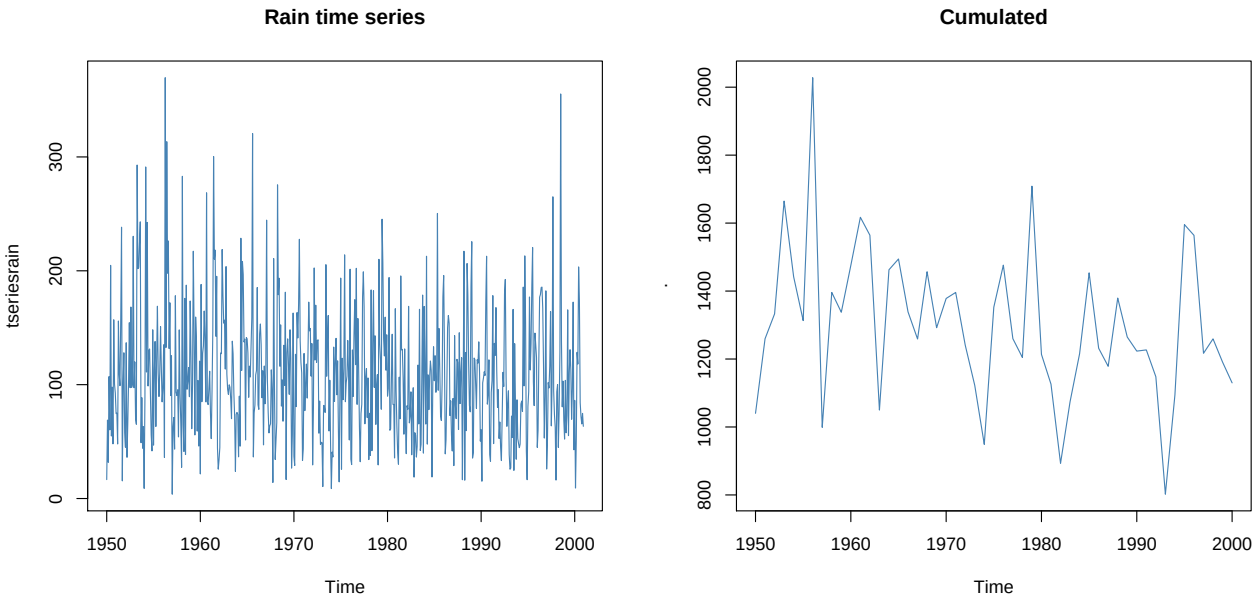
Question 1

```
raindata<-read.table('pluie1.txt',
                     sep='',
                     header = TRUE
                     )
head(raindata)
```

```
  annee mois pluie
1  1950    1  16.6
```

```
2  1950    2  68.9
3  1950    3  31.8
4  1950    4 107.1
5  1950    5  60.3
6  1950    6 204.8
```

```
tseriesrain<-ts(raindata$pluie,start = 1950,frequency = 12)
par(mfrow=c(1,2))
plot(tseriesrain,col='steel blue',main='Rain time series')
aggregate(tseriesrain,1,sum)%>%plot(col='steel blue',main='Cumulated')
```

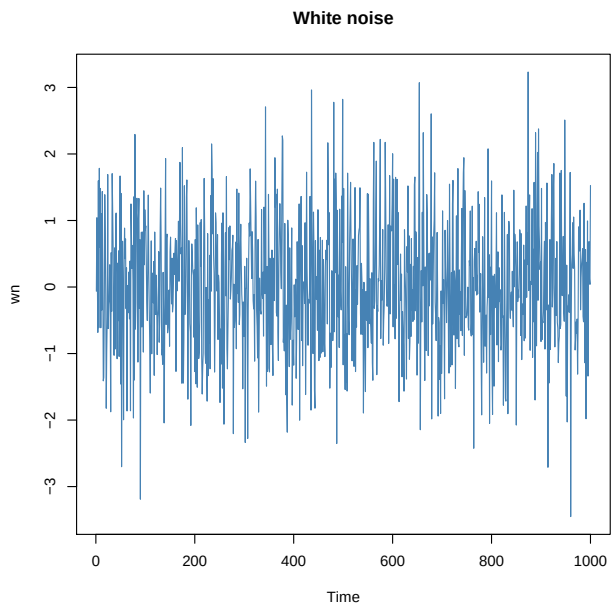


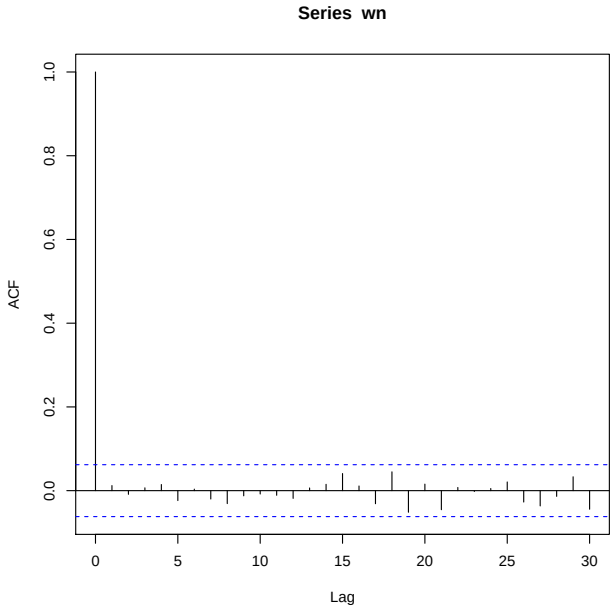
The last figure shows a downward trend. We also notice the independence of aggregated data.

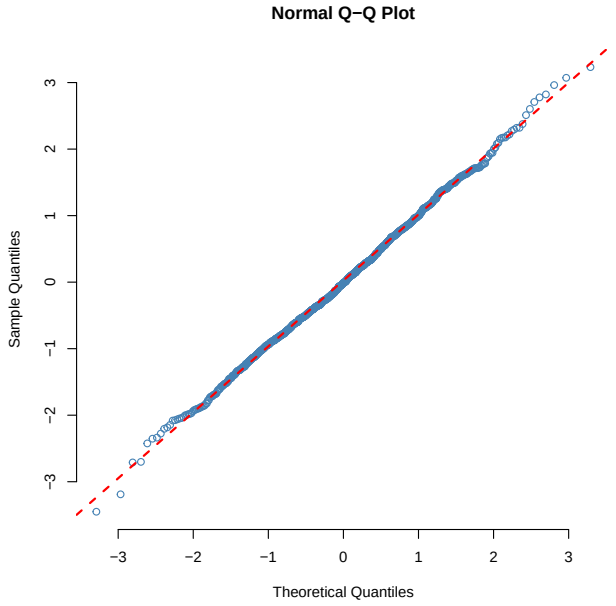
EXERCISE

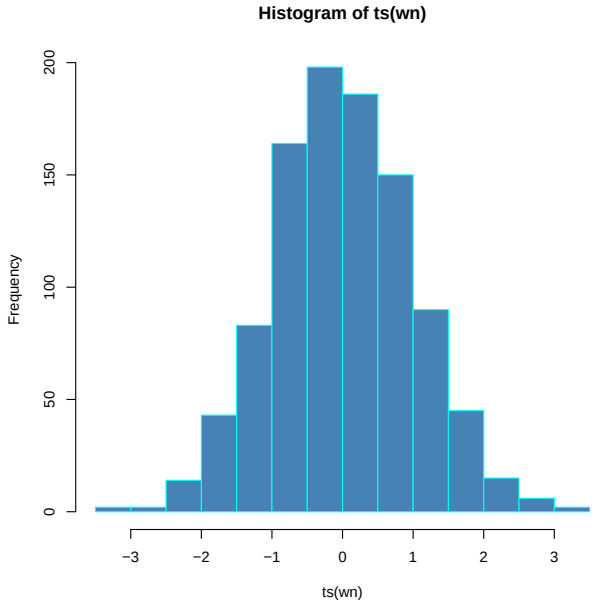
Question 1

```
n<-1000
wn<-rnorm(n)
par(mfrow=c(2,2))
plot.ts(wn,col='steel blue',main='White noise')
acf(wn)
qqnorm(ts(wn),
       col='steel blue',
       frame=FALSE,
       )
qqline(ts(wn),
       col='red',
       lty=2,
       lwd=2
       )
hist(ts(wn),col='steel blue',border = 'cyan')
```









Question 2

```
shapiro.test(ts(wn))
```

```
    Shapiro-Wilk normality test

data:  ts(wn)
W = 0.99875, p-value = 0.7188
```

The null hypothesis (observed data distribution is normal) can be reject.

Question 3

```
par(mfrow=c(2,3))
for(k in 1:6){rnorm(500)%>%ts()%>%acf()}
```