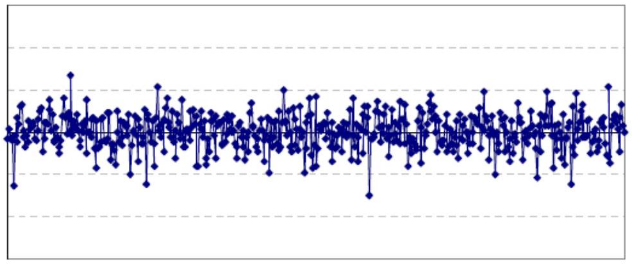
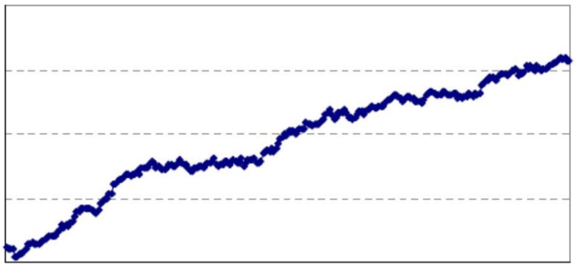
LECTURE 8 – TIME SERIES

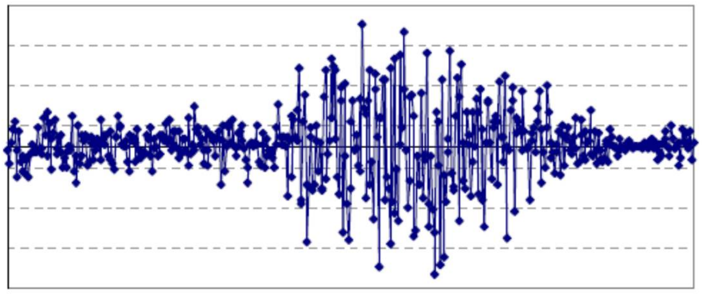
- Time series – definition
- Antoine Augustin Cournot (1801-1877) and William Stanley Jevons (1835- 1882) - data series observations
- Stationarity in time series analysis



(a) Serie de timp staționară în medie și varianță



(b) Serie de timp non-staționară în medie



(c) Serie de timp staționară în medie dar non-staționară în varianță

Fig. 1. Time series stationarity

There are two types of stationarity:

- *In mean (level)*
- *In variance*

Preliminary adjusting

- Box-Cox transform ($y_t = \frac{{x_t}^{\alpha}-1}{\alpha}$, $\alpha \neq 0$) of taking log
- Removing the seasonality
- Removing outliers
- Deflating and inflating the prices

Random walk processes

- 1900, Louise Jean-Baptiste Alphonse Bachelier – the thesis *Speculation theory* - approximated the prices evolution at the stock markets *by a random walk .*

Random walk process:

$$y_{t+1} = y_t + e_t,$$

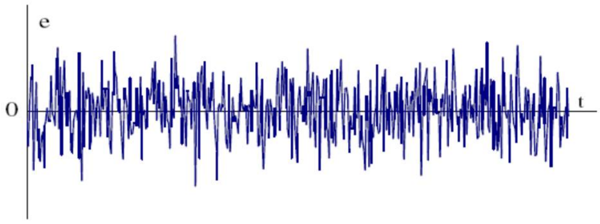


Fig. 2. Time series generated from N(0,1) – stationary in mean and variance

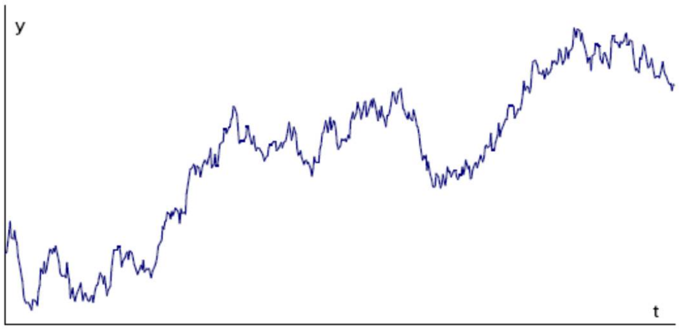


Fig. 3. Random walk process – nonstationary in mean

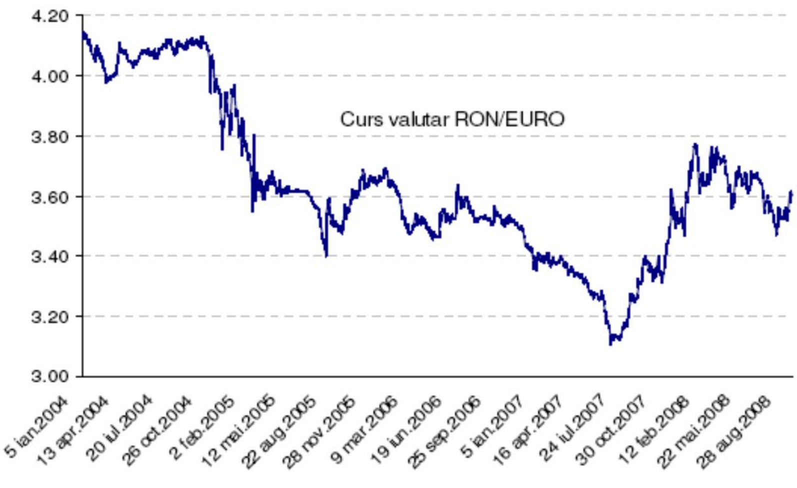


Fig. 4. Exchange rate Ron/Euro - Jan.2004-Sept. 2008 (it is not a random walk)

Deceomposition methods

The decomposition methods suppose the existence of three components: trend, seasonality and random variation.

The additive model is used in the case of a total independence of each component from the others, and the equation is:

$$Y = T + S + E,$$

where:

Y - the recorded variable;

T - the tendency;

S - seasonal variations;

E – the random variable.

The multiplicative model assumes the existence of a proportionality relation of the data series components, the observed values being the product of these components.

$Y = T \times S \times E$

The evolution of a time series with additive and multiplicative seasonality is presented in Fig. 4.

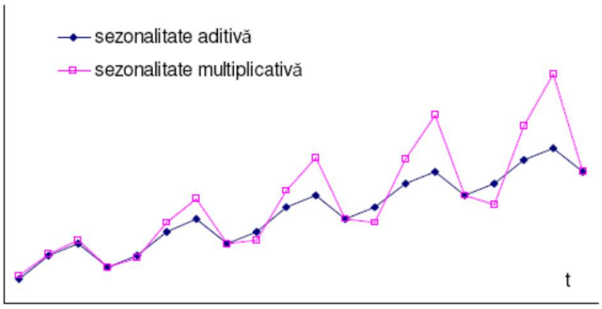


Fig.4. Types of seasonal variations

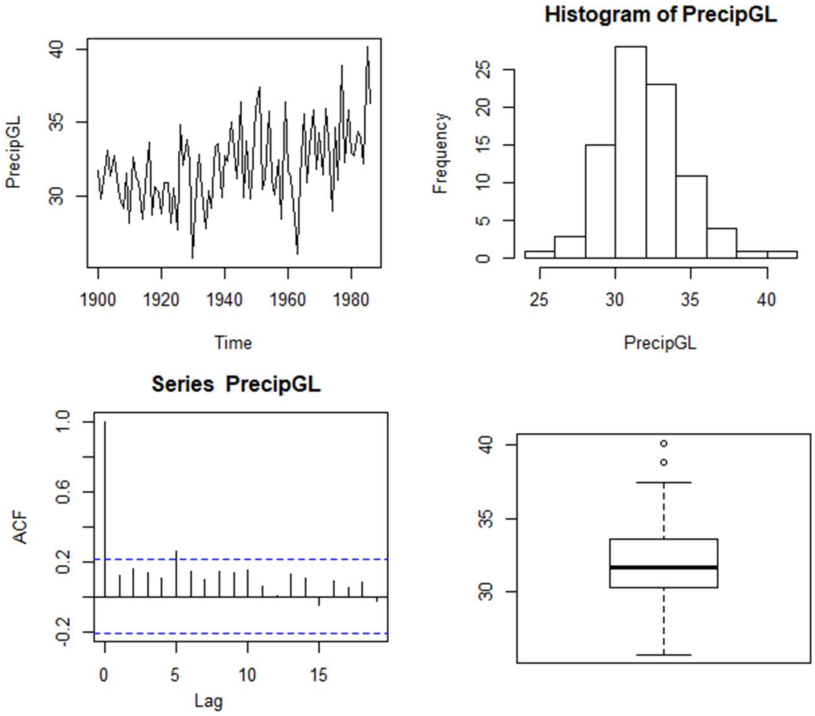
The trend component is determined by the moving average method of by analytical methods.

Time series analysis

1. **Graphical analysis**
   - Plot
   - Histogram – to remark the skewness
   - Boxplot and tests for the outliers detection – **outliers** package
   - autocorrelation function

2. **Statistical tests**
   - Homoskedasticity: the Bartlett test – Bartlett.test (see the previous laboratory)
   - Testing the hypotheses related to the trend existence – the Mann-Kendall test
     https://cran.r-project.org/web/packages/Kendall/Kendall.pdf
     https://cran.r-project.org/web/packages/trend/trend.pdf
     a. The null hypothesis is that there is no monotonic trend of the data series, while the alternative one is the existence of a monotonic trend

```
library(Kendall)
data(PrecipGL)
par(mfrow=c(2,2))
plot(PrecipGL)
hist(PrecipGL)
acf(PrecipGL)
boxplot(PrecipGL)
```



MannKendall(PrecipGL)

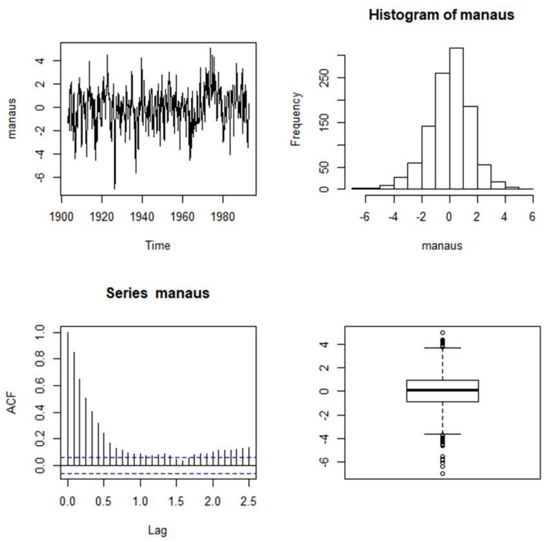
tau = 0.265, 2-sided pvalue =0.00029206 # monotonic trend

b. seasonal version of this test:

SeasonalMannKendall(PrecipGL)

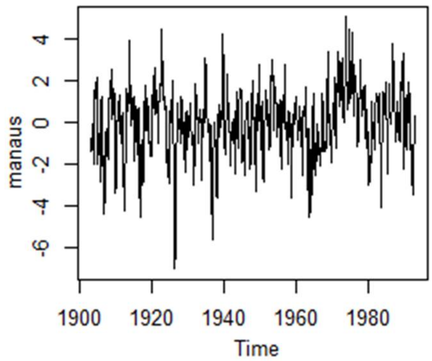
tau = 0.265, 2-sided pvalue =0.00028797 # trend sezonier monoton

library(boot)
data(manaus)
manaus



SeasonalMannKendall(manaus)

tau = 0.0877, 2-sided pvalue =2.2559e-05



Alternatively:

```
library(trend)
mk.test(manaus)
```

Mann-Kendall trend test

data: manaus
z = 4.3756, n = 1080, p-value = 1.211e-05
alternative hypothesis: true S is not equal to 0
sample estimates:

S	varS	tau
5.180400e+04	1.401620e+08	8.891681e-02

c. Determinarea tendintei liniare prin metoda neparametrica a lui **Sen** – **library trend**

This test computes both the slope (i.e. linear rate of change) and confidence levels according to Sen's method. First, a set of linear slopes is calculated as follows:

$$d_k = \frac{x_j - x_i}{j - i}$$

for $(1 \leq i < j \leq n)$, where d is the slope, x denotes the variable, n is the number of data, and i, j are indices.

Sen's slope is then calculated as the median from all slopes: $b_{Sen} = \text{median}(d_k)$.

This function also computes the upper and lower confidence limits for sens slope.

```
library(trend)
sens.slope(PrecipGL)
```

Sen's slope

data: PrecipGL
z = 3.6222, n = 87, p-value = 0.0002921
alternative hypothesis: true z is not equal to 0
95 percent confidence interval: [0.01952381, 0.06217391]
sample estimates: Sen's slope: 0.04

```
sens.slope(manaus)
```

Sen's slope
data: manaus

z = 4.3756, n = 1080, p-value = 1.211e-05
alternative hypothesis: true z is not equal to 0
95 percent confidence interval: [0.0003428378, 0.0009065682]
sample estimates: Sen's slope: 0.0006241918

- Testing the stationarity

  \- in level and trend – KPSS test: the null hypothesis is the stationarity in level (trend), while the opposite one is the non-stationarity in level (trend)

  https://cran.r-project.org/web/packages/tseries/tseries.pdf

```
library(tseries)
kpss.test(x, null = c("Level", "Trend"), lshort = TRUE)

kpss.test(PrecipGL)

        KPSS Test for Level Stationarity

    data:  PrecipGL
    KPSS Level = 1.1118, Truncation lag parameter = 3, p-value = 0.01

    Warning message: In kpss.test(PrecipGL) : p-value smaller than printed p-value

kpss.test(PrecipGL, null = c("Trend"))

           KPSS Test for Trend Stationarity

       data:  PrecipGL
       KPSS Trend = 0.091423, Truncation lag parameter = 3, p-value = 0.1

       Warning message:
       In kpss.test(PrecipGL, null = c("Trend")) :  p-value greater than printed p-value
    ## stationarity in trend and non-stationarity in level
```

  \- Dickey-Fuller and ADF (Augmented Dickey-Fuller) test, where the null hypothesis is the existence of an unit root, and the alternative one is the stationarity

  https://cran.r-project.org/web/packages/tseries/tseries.pdf

```
library(tseries)

       adf.test(PrecipGL)

        Augmented Dickey-Fuller Test

    data:  PrecipGL
    Dickey-Fuller = -3.4123, Lag order = 4, p-value = 0.05846 #there is an unit root
    alternative hypothesis: stationary
##The series is not stationary
```

- Detecting the change points – Pettitt, Buishand U Test, CUSUM – in package **trend.**

```
pettitt.test(PrecipGL)

        Pettitt's test for single change-point detection

data:  PrecipGL
U* = 929, p-value = 0.0008408  # there is a change point
alternative hypothesis: two.sided
sample estimates:
probable change point at time K    37

out <- bu.test(PrecipGL)

out

        Buishand U test

data:  PrecipGL
U = 1.5316, n = 87, p-value = 5e-05
alternative hypothesis: true delta is not equal to 0
sample estimates:
probable change point at time K  37
```

- Packages for multiple change points detection
  https://cran.r-project.org/web/packages/changepoint/changepoint.pdf