Time Series
1.Stochastic Processes

Josep A. Sanchez-Espigares

Department of Statistics and Operations Research
Universitat Politècnica de Catalunya - BarcelonaTECH
Barcelona, Spain

josep.a.sanchez@upc.edu

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Outline

1. Introduction to time series
2. Stationary series
3. ACF, PACF

Bibliography

Textbook

- Shumway R., Stoffer D. (2016). *Time Series Analysis and Its Applications - With R Examples*.

http://www.stat.pitt.edu/stoffer/tsa4/tsa4.htm

References

- Box G., Jenkins G., Reinsel G. (2008). *Time series Analysis: forecasting and control*
- Peña D. (2005). *Análisis de Series Temporales*

Time series: Ordered sequence of observations of the same phenomenon. Tipically measured at equally spaced successive instants of time.

$$\{X_t\}_{t=1,\ldots,T} = \{X_1, X_2, \ldots X_T\}$$

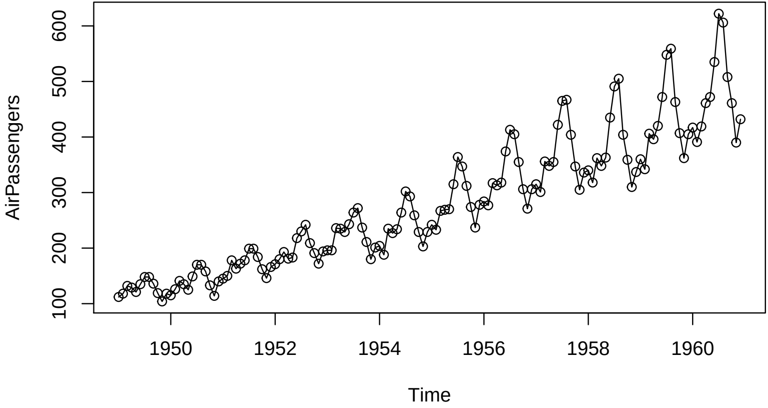
Example:

`AirPassengers`: Monthly totals of international airline passengers in USA, 1949 to 1960 (Box & Jenkins, 1976)

```
##      Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
## 1949 112 118 132 129 121 135 148 148 136 119 104 118
## 1950 115 126 141 135 125 149 170 170 158 133 114 140
## 1951 145 150 178 163 172 178 199 199 184 162 146 166
## 1952 171 180 193 181 183 218 230 242 209 191 172 194
## 1953 196 196 236 235 229 243 264 272 237 211 180 201
## 1954 204 188 235 227 234 264 302 293 259 229 203 229
## 1955 242 233 267 269 270 315 364 347 312 274 237 278
## 1956 284 277 317 313 318 374 413 405 355 306 271 306
## 1957 315 301 356 348 355 422 465 467 404 347 305 336
## 1958 340 318 362 348 363 435 491 505 404 359 310 337
## 1959 360 342 406 396 420 472 548 559 463 407 362 405
## 1960 417 391 419 461 472 535 622 606 508 461 390 432
```

Time series definition

```
plot(AirPassengers,type="o")
```



Motivation and Objetives

Motivation

- Describing and forecasting time series is crucial in different areas of knowledge; including finance, econometrics, signal processing and a long etc.

Objectives

- **Description**: Describe temporal patterns in a time series: regular and/or seasonal effects, cyclicity, trends, outliers, sudden changes, breaks, $\cdots$
- **Estimation**: Estimate the values of the time series parameters
- **Validation**: Validate the estimated parameters and decide if the estimated parameters are significant or not.
- **Prediction/Forecasting**: Predict future values of the time series.

Plot of the series and identification of the components:

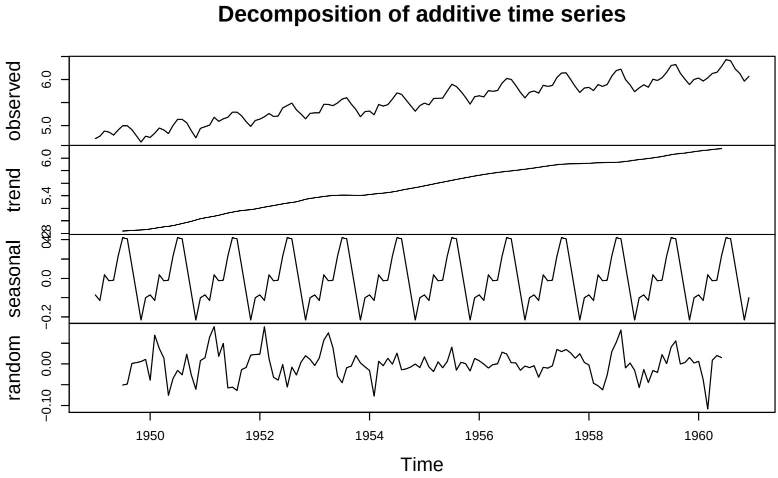
- Trend(T_t): Long term tendency
  - Moving average of order s:

$$T_t = \frac{1}{s}\sum_{i=1}^{s} X_{t-s/2+i}$$

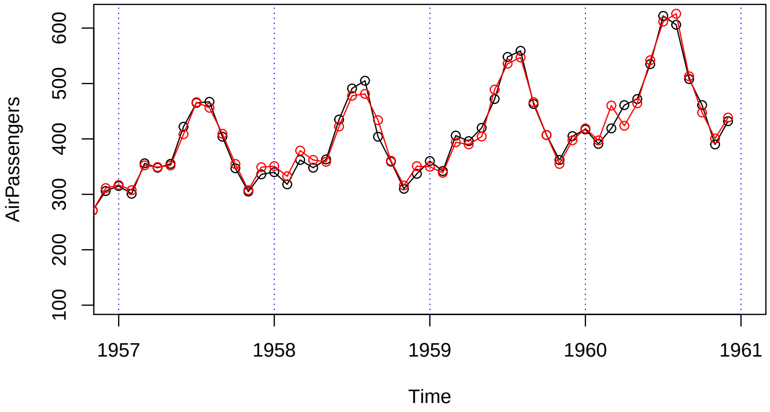
- Seasonal(S_t): Pattern repeated periodically with the same period
  - Seasonal index: Mean for each period of detrended series $(X_t - T_t)$
- Cycle(C_t): Pattern repeated periodically with non-constant period
  - Not easy to model due to the changing period
- Random(w_t): Random noise
  - Remainder $(X_t - T_t - S_t - C_t)$

Additive model:

$$X_t = T_t + S_t + C_t + w_t$$


Decomposition of additive time series
observed
trend
seasonal
random
6.0
5.0
6.0
5.4
4.8
0.2
0.0
−0.2
0.00
−0.10
1950
1952
1954
1956
1958
1960
Time

Goal: Find a mathematical model that reflects the behaviour of the observed data



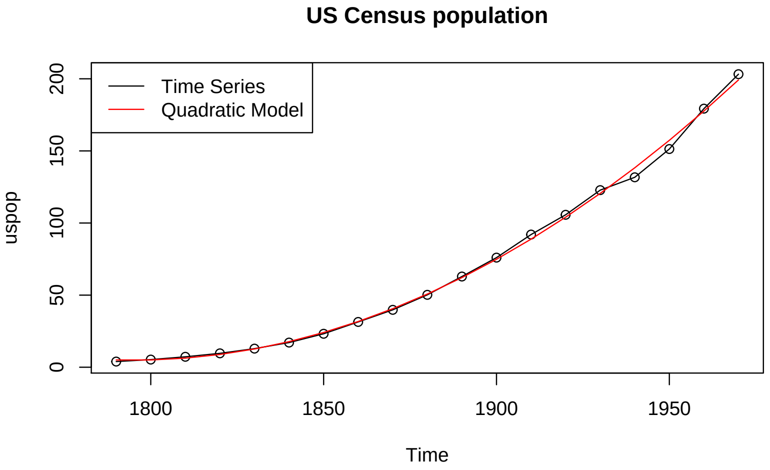
- **Deterministic model**: The expected value of X_t depends on a parametric function F of t and the random component does not depend on the previous values.

$$X_t = F(t) + Z_t \quad Z_t \sim N(0, \sigma_Z^2)$$

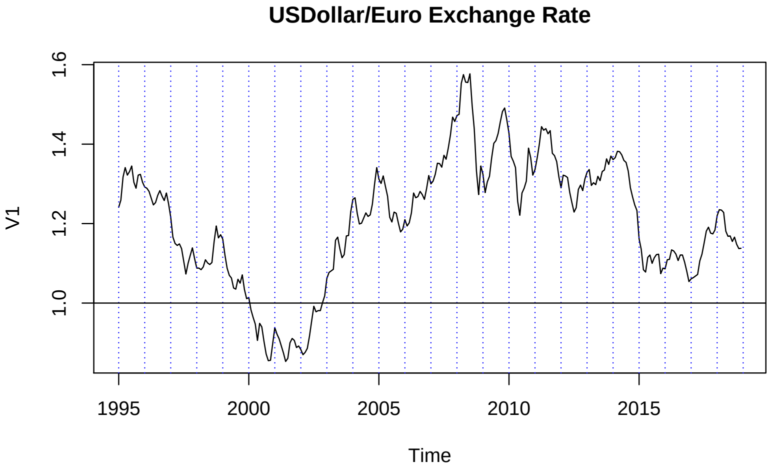
- **Stochastic Model**: The expected value of X_t depends on the previous values $X_{t-1}, X_{t-2}, ...$ and/or the previous random components $Z_{t-1}, Z_{t-2}, ...$ plus a random component independent of the past.

$$X_t = G(X_{t-1}, X_{t-2}, ..., Z_{t-1}, Z_{t-2}, ...) + Z_t \quad Z_t \sim N(0, \sigma_Z^2)$$

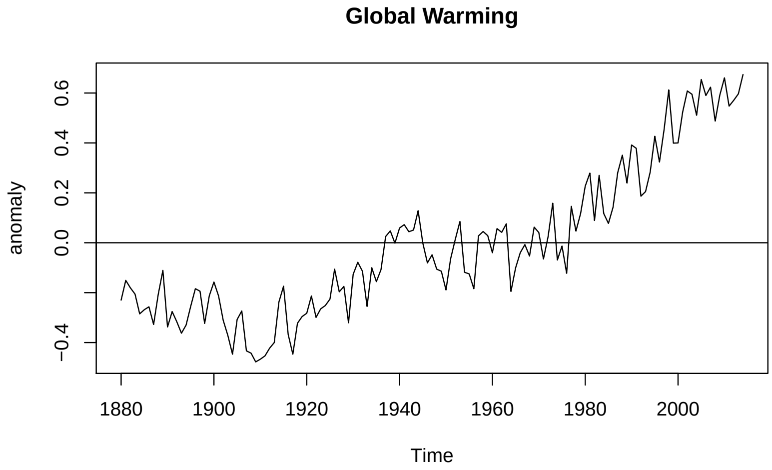
Example 1.1: Population recorded by US Census, 19 decades, 1790 to 1970.



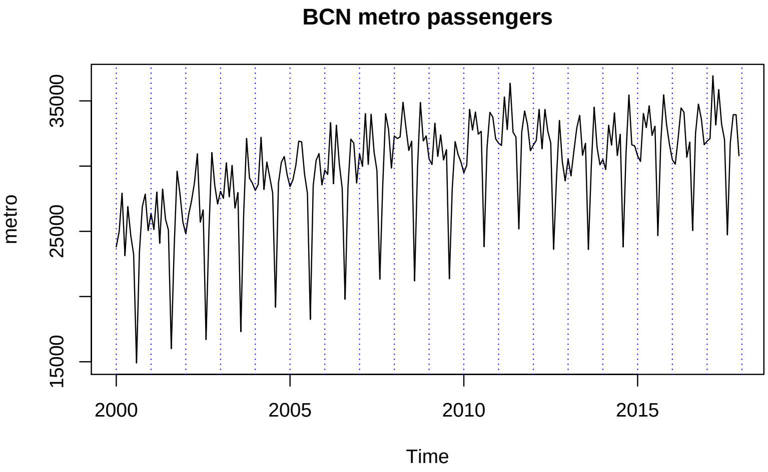
Example 1.2: Exchange Rate Dollar/Euro (ECU before 1999). Monthly mean. Source: Bank of Spain



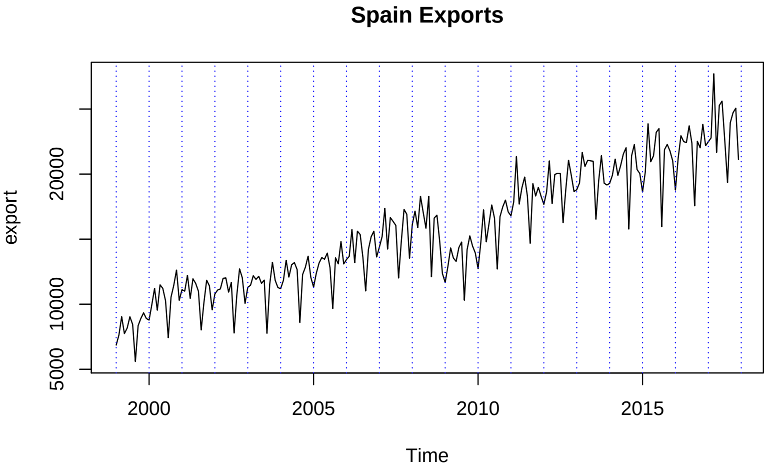
Example 1.3: Global Warming: Yearly average global temperature deviations (1880-2009) in degrees centigrade. Source: NASA



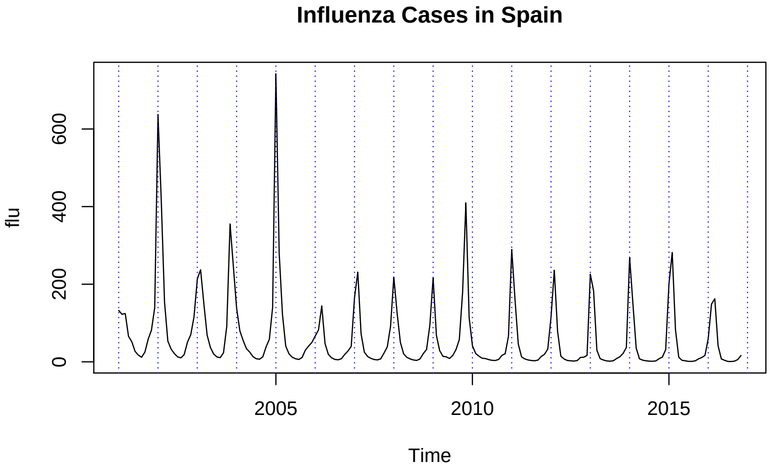
Example 1.4: Barcelona metro passengers (thousands). Monthly data. Source: INE



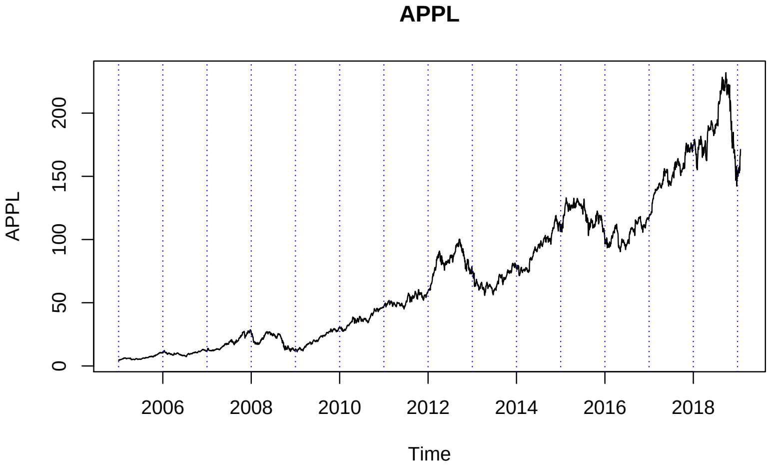
Example 1.5: Spain: Total Exports (thousand of millions). Source: Ministry of industry, trade and tourism of Spain



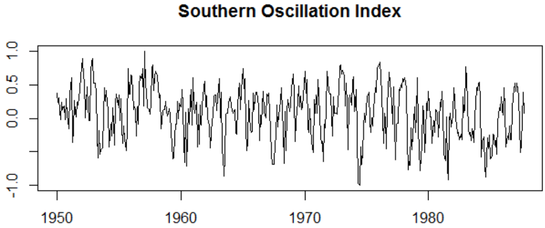
Example 1.6: Number of reported cases of influenza affected (thousand). Monthly data. Source: Ministry of Health of Spain

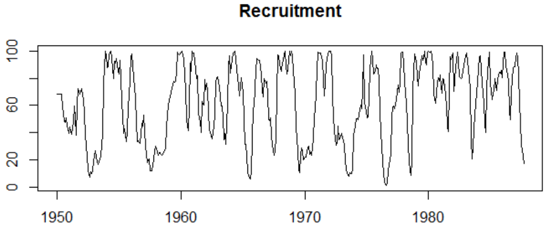


Example 1.7: Apple Inc.(AAPL) NasdaqGS Real Time Price. Currency in USD

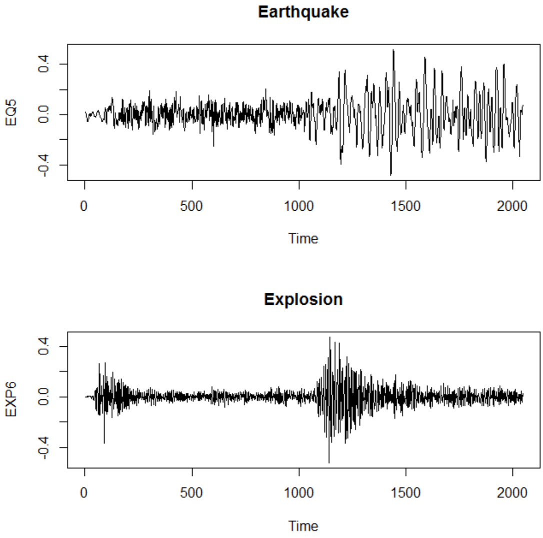


Example 1.8: El Nino and Fish Population. Monthly Southern Oscillation Index (SOI) and Recruitment (estimated new fish), 1950-1987.

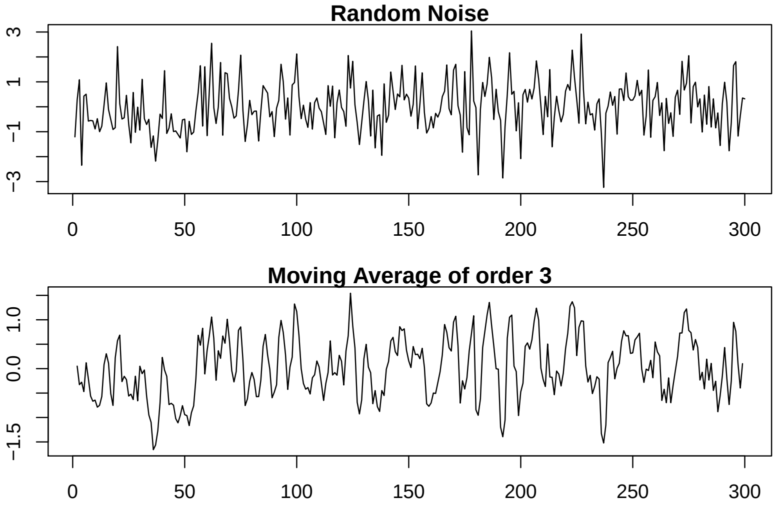




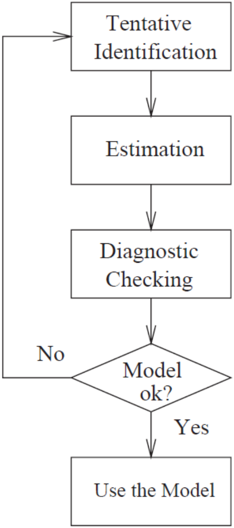
Example 1.9: Earthquakes and Explosions (Arrival phases from an earthquake (top) and explosion (bottom) at 40 points per second.



Example 1.10: Gaussian white noise series (top) and three-point moving average of the Gaussian white noise series (bottom).



Box-Jenkins Methodology



Time Series Plot
Range-Mean Plot
ACF and PACF

Least Squares or
Maximum Likelihood

Residual Analysis
and Forecasts

Forecasting
Explanation
Control

Distribution of a general stochastic process

- First and second moments for the multivariate distribution of $\{X_t\}_{t=1..T}$

$$E[(X_1, X_2, ...X_T)] = (\mu_1, \mu_2, ..., \mu_T)$$

$$Var((X_1, X_2, ...X_T)) = \begin{pmatrix} \sigma_1^2 & \sigma_{1,2} & \sigma_{1,3} & \cdots & \sigma_{1,T} \\ \sigma_{1,2} & \sigma_2^2 & \sigma_{2,3} & \cdots & \sigma_{2,T} \\ \sigma_{1,3} & \sigma_{2,3} & \sigma_3^2 & \cdots & \sigma_{3,T} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{1,T} & \sigma_{2,T} & \sigma_{3,T} & \cdots & \sigma_T^2 \end{pmatrix}$$

- Parameters of the model
  - T values for the mean: $E(X_t) = \mu_t$
  - T values for the variances: $V(X_t) = \sigma_t^2$
  - $T*(T-1)$ values for the covariances: $Cov(X_t, X_s) = \sigma_{t,s}$

Distribution of an stationary stochastic process

- First and second moments for the multivariate distribution of $\{X_t\}_{t=1..T}$

$$E[(X_1, X_2, ...X_T)] = (\mu, \mu, ..., \mu)$$

$$Var((X_1, X_2, ...X_T)) = \begin{pmatrix} \sigma^2 & \sigma_1 & \sigma_2 & \dots & \sigma_{T-1} \\ \sigma_1 & \sigma^2 & \sigma_1 & \dots & \sigma_{T-2} \\ \sigma_2 & \sigma_1 & \sigma^2 & \dots & \sigma_{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{T-1} & \sigma_{T-2} & \sigma_{T-3} & \dots & \sigma^2 \end{pmatrix}$$

- Parameters of the model
  - 1 value for the mean: $E(X_t) = \mu$
  - 1 value for the variances: $V(X_t) = \sigma^2$
  - $T-1$ values for the covariances: $Cov(X_t, X_s) = \sigma_{|t-s|}$

- **Strict Stationary process** or series has the following **properties**:
  - the joint distribution of the whole series does not depend on the time origin

$$F_{(X_1,\dots,X_t)}(x_1,\dots,x_t) = F_{(X_{1+s},\dots,X_{t+s})}(x_{1+s},\dots,x_{t+s}) \quad \forall t, s$$

- **Weakly Stationary process** or series has the following **properties**:
  - the two first moments of the multivariate distribution of the whole series does not depend on the time origin:
    - constant mean (μ)
    - constant variance (σ^2)
    - constant autocovariance structure ($\sigma_{t,s} = \sigma_{|t-s|}$)
    - The latter refers to the covariance between X_t and X_{t-1} being the same as X_{t-s} and X_{t-s-1}.

Weakly Stationary Process + Gaussian multivariate Distribution

$\Rightarrow$ Strict Stationary Process

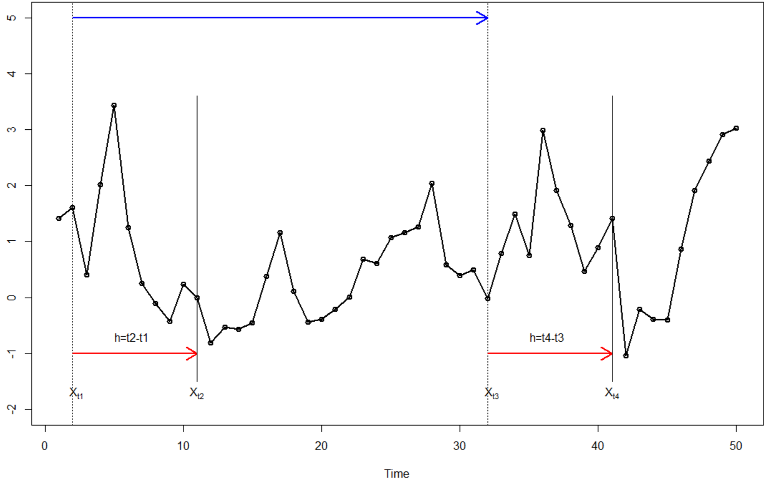
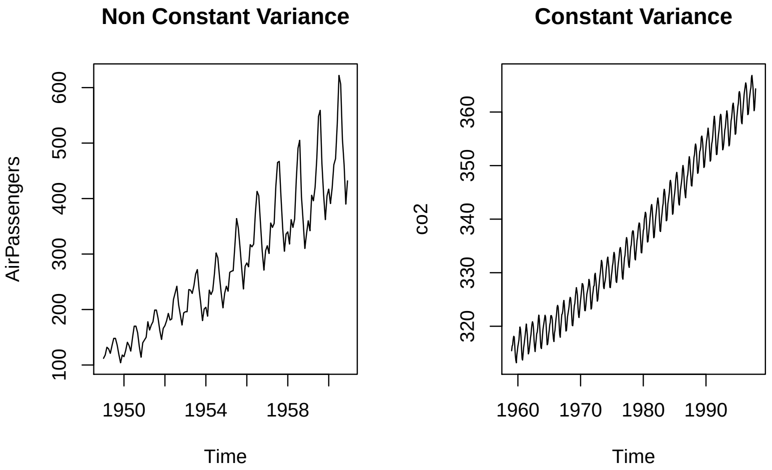


Figure 1: Example of an stationary process

- Is our data stationary?
- How can we detect?
- In general:
  - **Plot** the data
  - Identify no stationary components (trends, seasonal patterns, cycles)
  - Transform the series to remove those components
  - For the transformed (stationary) series, plot and analyze the sample autocorrelation

Is the variance constant?

It is very common that the variance of the series increases when the level of the series rises:



Transformations: Change the Scale

Tools to diagnose the non-constant variance:

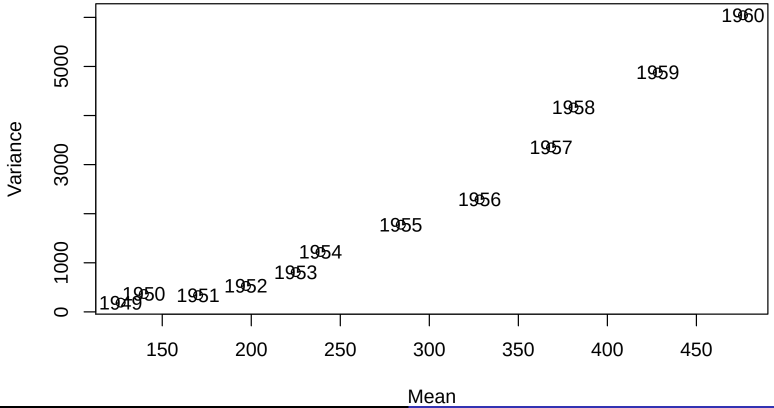
1. Mean-Variance plot:
   - Calculate the mean and the variance of consecutive groups of 8-12 observations
   - Plot the variance against the mean of each group
2. Boxplot for periods:
   - Represent the boxplot for each group of 8-12 observations
   - The height of the box (IQR) is a robust estimate of variability

- If the variance is similar for all the groups $\Rightarrow$ No scale transformation
- If the variance is higher for higher values of the mean $\Rightarrow$ Change the scale
  - Box-Cox transformation:

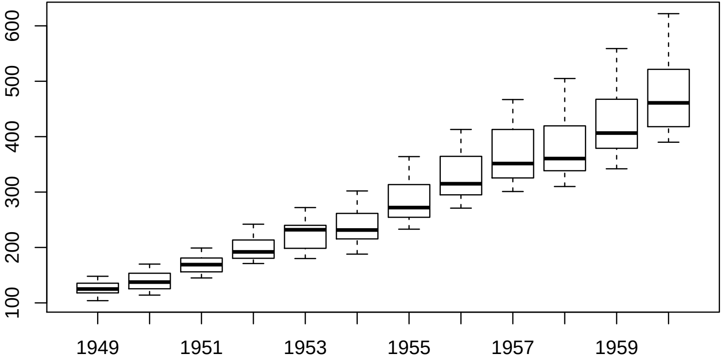
$$\begin{cases} \frac{X^{\lambda}-1}{\lambda} & \lambda \in [-1,2], \lambda \neq 0 \\ \log(X) & \lambda = 0 \end{cases}$$

- Note: Usually the log transformation is applied (easy to interpret)

```r
m=by(AirPassengers,floor(time(AirPassengers)),mean)
v=by(AirPassengers,floor(time(AirPassengers)),var)
plot(v~m,xlab="Mean",ylab="Variance")
text(m,v,1949:1960)
```



```
boxplot(AirPassengers~floor(time(AirPassengers)))
```



Is there a Seasonal Pattern?

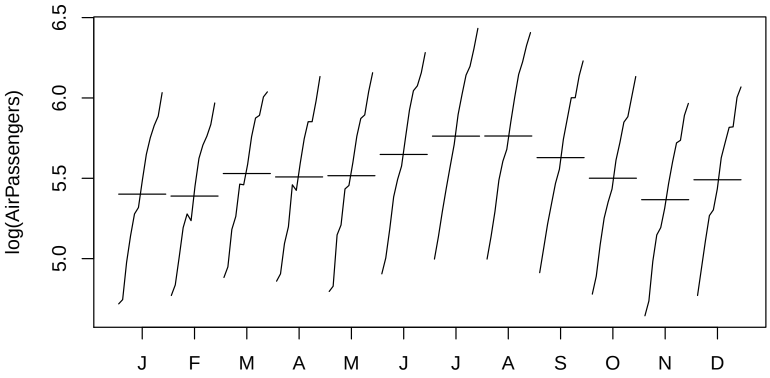
- A similar pattern for a constant period **s** is observed
  - Monthly data: s=12 observations
  - Quarterly data: s=4 observations
  - Daily data: s=7 observations
  - Hourly data: s=24 observations
- This pattern is the so-called **Seasonal Pattern** of the time series.
- To remove this pattern, a linear filter is applied to the series
  - Moving Average of order s: $W_t = \frac{1}{s}\sum_{i=1}^{s} X_{t-i+1}$
  - Seasonal difference of order s: $W_t = X_t - X_{t-s} \quad t > s$

Note: The seasonal difference is prefered and includes the other option

Transformation: Seasonal difference

Tool to diagnose the seasonal pattern:

```
monthplot(log(AirPassengers))
```



Transformation: Seasonal difference

Notation Backshift operator: $BX_t = X_{t-1}$ $\quad B^s X_t = X_{t-s}$

(same as lag operator **L** in some articles/books)

Algebraic notation:
- Moving Average of order s:

$$W_t = \frac{1}{s}\sum_{i=1}^{s} X_{t-i+1} = (1 + B + ... + B^{s-1})X_t$$

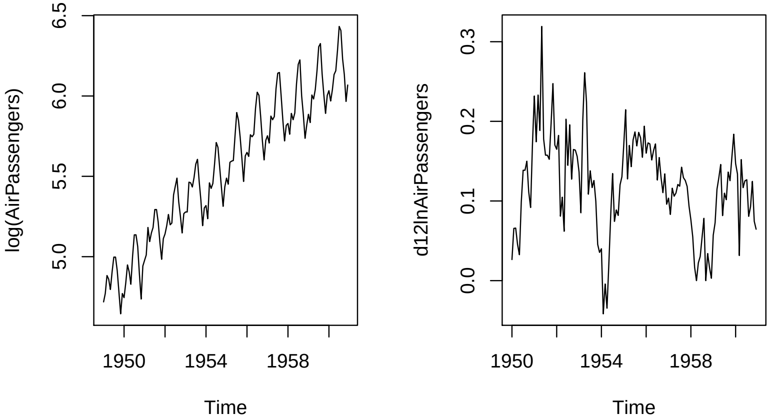
- Seasonal Difference of order s:

$$W_t = X_t - X_{t-s} = (1 - B^s)X_t = \nabla_s X_t$$

Note: The seasonal difference is equivalent to a regular difference of a moving average of order s

$$(1 - B^s) = (1 - B)(1 + B + ... + B^{s-1})$$

Transformation: Seasonal difference



Transformation: Regular difference

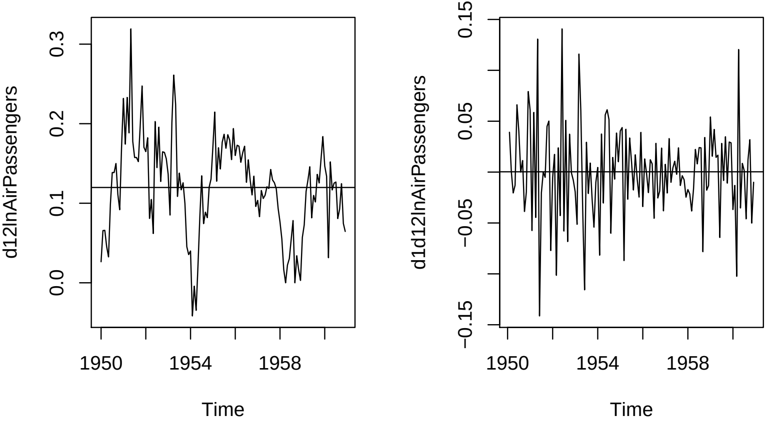
Is the mean constant?

- Linear or general trend implies non constant mean.
- A regular difference is applied until the mean can be considered constant

$$W_t = X_t - X_{t-1} = (1 - B)X_t$$

- **Overdiferentation**: If the differenced time series yields a higher variance, then the later difference is not needed

Transformation: Regular difference



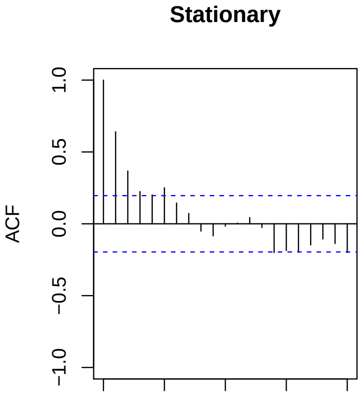
Notation Backshift operator: $BX_t = X_{t-1}$ $\quad B^s X_t = X_{t-s}$

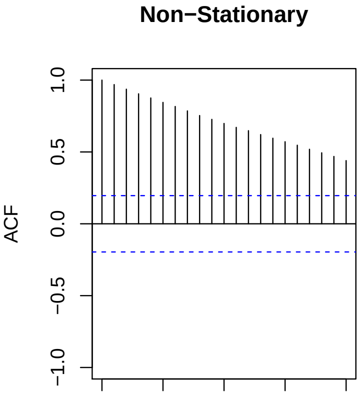
(same as lag operator **L** in some articles/books)

NON-STATIONARITY CAUSE	TRANSFORMATION
Non-constant Variance	Box-Cox Transformation ($\lambda \in [-1, 2]$ $W_t = \frac{x_t^\lambda - 1}{\lambda} \quad \lambda \neq 0$ $W_t = \log X_t \quad \lambda = 0$
Linear Deterministic Trend Non-constant mean	Regular difference $W_t = (1 - B)X_t$ Regular difference $W_t = (1 - B)X_t$
Deterministic d-order polynomial Trend Stochastic Trend	d-Regular differences $W_t = (1 - B)^d X_t$ d-Regular differences $W_t = (1 - B)^d X_t$ until stationary W_t
Seasonal pattern of order s	Seasonal difference $W_t = (1 - B^s)X_t$
Indexes and Financial data	log-Returns: $W_t = (1 - B) \log X_t \cong \frac{X_t - X_{t-1}}{X_{t-1}}$

MOMENT	THEORETICAL	SAMPLE
Mean	μ	$\bar{X}_t = \frac{1}{T}\sum_{t=1}^{T} X_t$
Autocovariance $\gamma(k)$	$E\left[(X_{t+k}-\mu)(X_t-\mu)\right]$	$\frac{1}{T}\sum_{t=1}^{T-k}(X_{t+k}-\bar{X})(X_t-\bar{X})$
Variance $\sigma_X^2 = \gamma(0)$	$E\left[(X_t-\mu)^2)\right]$	$\frac{1}{T}\sum_{t=1}^{T}(X_t-\bar{X})^2$
Autocorrelation $\rho(k) = \gamma(k)/\gamma(0)$	$\frac{E\left[(X_{t+k}-\mu)(X_t-\mu)\right]}{E\left[(X_t-\mu)^2)\right]}$	$\frac{\sum_{t=1}^{T-k}(X_{t+k}-\bar{X})(X_t-\bar{X})}{\sum_{t=1}^{T}(X_t-\bar{X})^2}$

- **Autocorrelation Function (ACF)**: measures the relationship between the two k-lag apart variables, X_t and X_{t+k} .
- **ACF** lies between -1 and +1
- Correlogram is the plot of the ACF $\rho(k)$ against k
- Under **Stationarity**: ACF falls immediately from 1 to 0
- Under **Non-stationary**: the ACF declines gradually from 1 to 0 over a prolonged period of time





ACF, PACF: Correlogram

Variance of the sample ACF:

- For large sample size T, asymptotically:

$$V(\hat{\rho}(k)) \approx \frac{1}{T}$$

The sample ACF represents the values of $\hat{\rho}(k)$ for each lag k from $k = 1, 2, \ldots$ The confidence bands are calculated using the asymptotic distribution for the estimator:

$$\pm\frac{1.96}{\sqrt{T}}$$

For each lag k we can test its significance by using the plot:

- If $\hat{\rho}(k)$ lies between the confidence bands, we cannot reject the null hypothesis $(H_0 : \rho(k) = 0)$ and the theoric autocorrelaction for this lag can be considered null.

ACF, PACF: Partial ACF

PACF: Partial correlation (of a stationary process) is the relationship between two variables, after excluding the effect of one or more independent variables.

In other words:

- $\phi_{11} = cor(X_{t+1}, X_t) = \rho(1)$
- $\phi_{hh} = cor(X_{t+h} - \hat{X}_{t+h}, X_t - \hat{X}_t),\ h \geq 2$
- Partial Autocorrelation Function (PACF) is similar to the ACF
- For instance, consider a regression context in which y = response variable and x_1, x_2, and x_3 are predictor variables. The **partial correlation** between y and x_3 is the correlation between the variables determined taking into account how both y and x_3 are related to x_1 and x_2

Partial Autocorrelation function

Ordinary Least Squares:

$$x_t = \phi_{\mathbf{1},\mathbf{1}} x_{t-1} + Z_t$$

$$x_t = \phi_{1,2} x_{t-1} + \phi_{\mathbf{2},\mathbf{2}} x_{t-2} + Z_t$$

$$x_t = \phi_{1,3} x_{t-1} + \phi_{2,3} x_{t-2} + \phi_{\mathbf{3},\mathbf{3}} x_{t-3} + Z_t$$

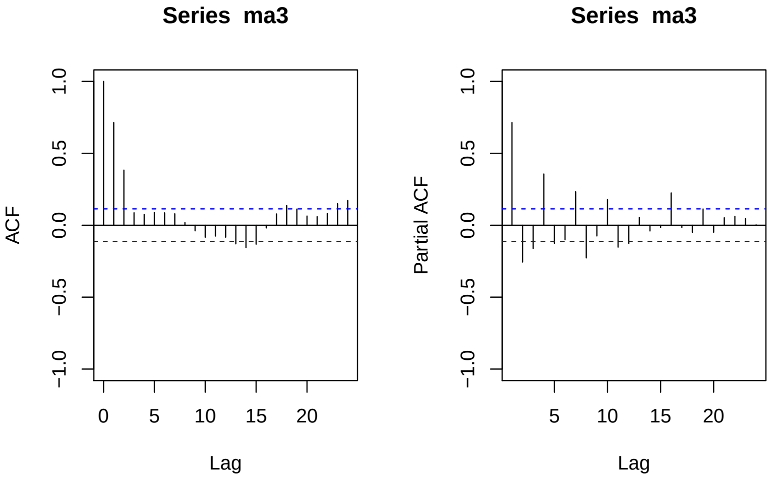
$$\vdots$$

$$x_t = \phi_{1,h} x_{t-1} + \phi_{2,h} x_{t-2} + \phi_{3,h} x_{t-3} + \ldots + \phi_{\mathbf{h},\mathbf{h}} x_{t-h} + Z_t$$

$$\vdots$$

PACF: $\{\phi_{\mathbf{1},\mathbf{1}}, \phi_{\mathbf{2},\mathbf{2}}, \ldots, \phi_{\mathbf{h},\mathbf{h}}, \ldots\}$

Sample **ACFs** and **PACF**



Standard

R: Sample ACF begins at 0 but sample PACF begins at 1

*Autocorrelation of white noise (Z_t)

$$Z_t \sim WN(\sigma_Z^2)$$
$$\sim N(0, \sigma_Z^2)$$

Independent

MOMENT	THEORETICAL
Mean	0
Autocovariance $\gamma(k)$	0
Variance $\gamma(0)$	σ_Z^2
Autocorrelation $\rho(k) = \gamma(k)/\gamma(0)$	0

Sample **ACF** and **PACF** for a white noise series