Stochastic processes I

The ASTA team

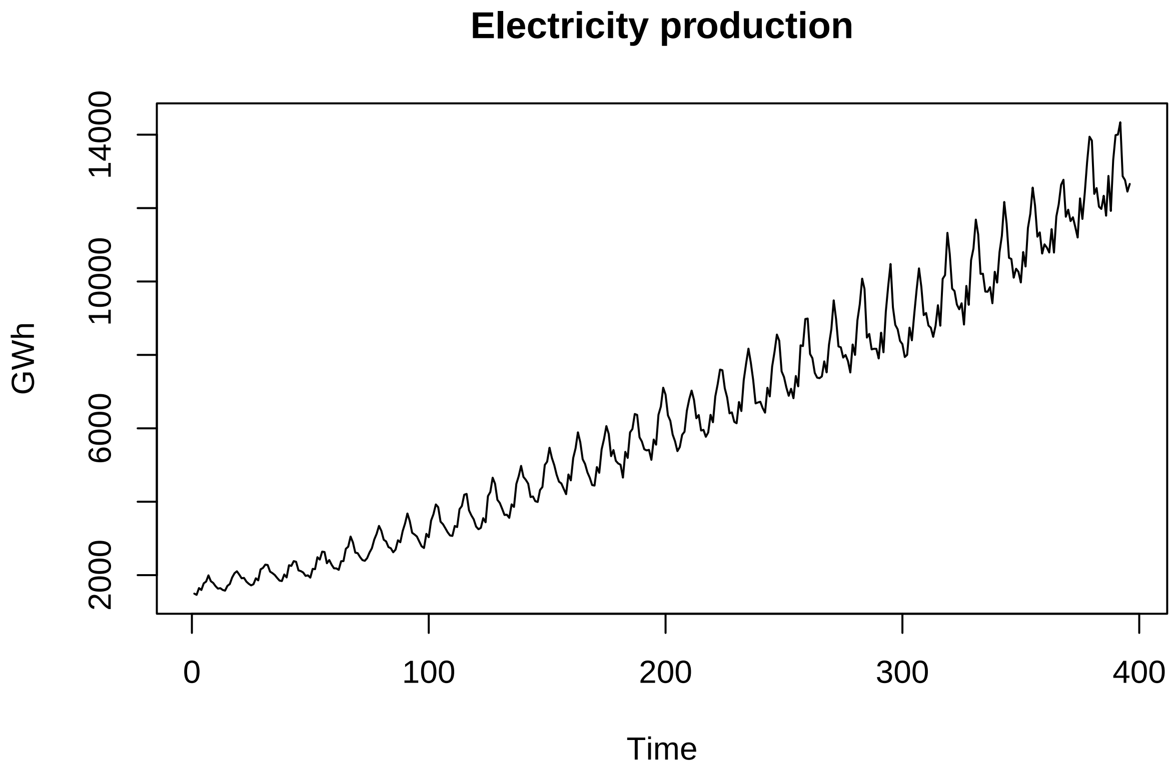
Contents

1 Introduction to stochastic processes

1.1 Data examples

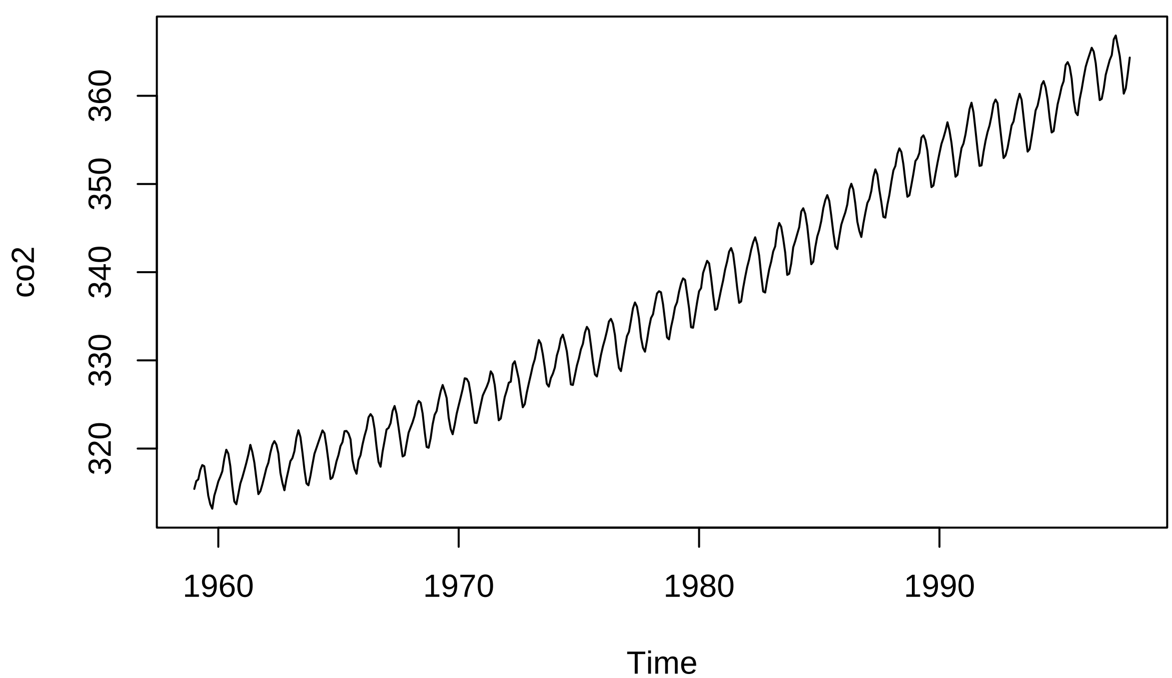
- A special type of data arises when we measure the same variable at different points in time with equal steps between time points.
- This data type is called a (discrete time) **stochastic process** or a **time series**
- One example is the time series of monthly electricity production (GWh) in Australia from Jan. 1958 to Dec. 1990 :

```
CBEdata <- read.table("https://asta.math.aau.dk/eng/static/datasets?file=cbe.dat", header = TRUE)
CBE <- ts(CBEdata[,3])
plot(CBE, ylab="GWh",main="Electricity production")
```



- Another example is monthly measurements of the atmospheric CO_2 concentration measured at Mauna Loa 1959 - 1997:

```
dat<-ts(co2)
plot(co2)
```



- Other examples:
  - Hourly wind speed measurements
  - Daily elspot prices

    - An electrical signal measured each millisecond
- Aim: Model, analyse and make predictions for such datasets.

1.2 Stochastic processes

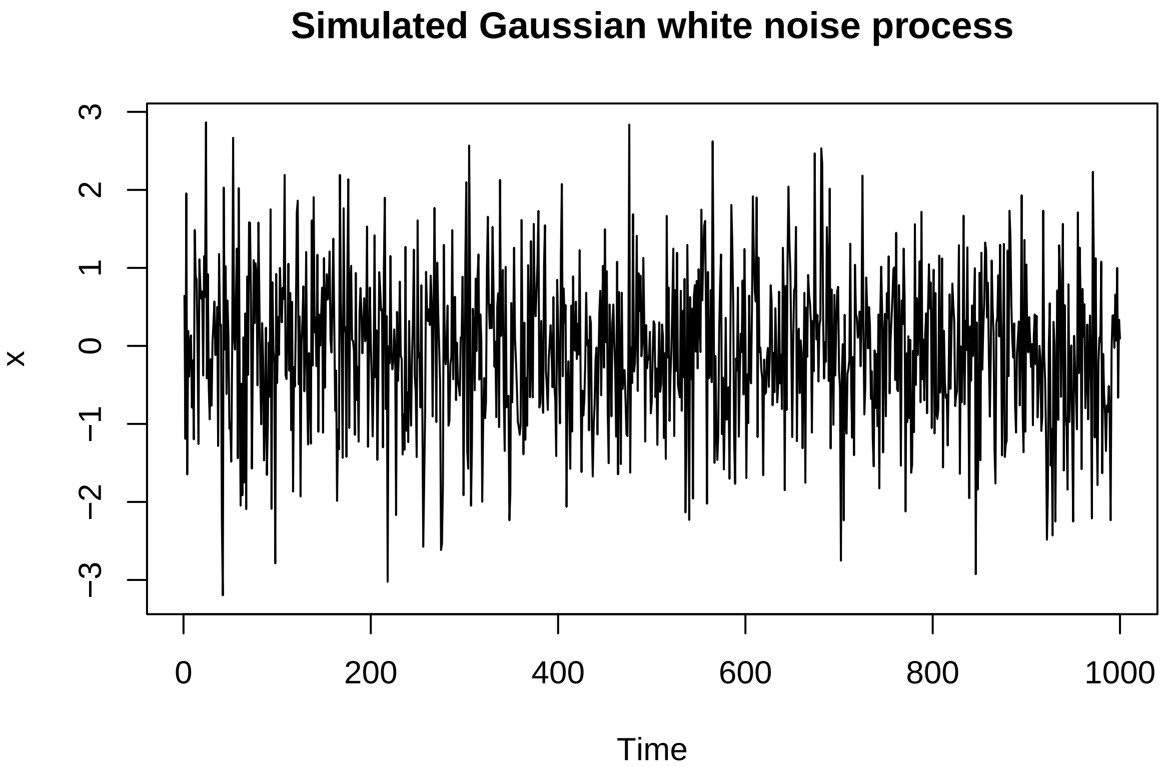
- We denote by X_t the variable at time t. We denote the time points by $t = 1, 2, 3, \ldots, n$.
    - We will always assume the data is observed at equidistant points in time (i.e. time steps between consecutive observations are the same).
- Measurements that are close in time will typically be similar: observations are not statistically independent!
- Measurements that are far apart in time will typically be less correlated.

2 Important stochastic processes

2.1 Example 1: White noise

- A stochastic process is called a **white noise process** if the X_t are
    - statistically independent
    - identically distributed
    - have mean 0 and variance σ^2
- It is called **Gaussian white noise**, if
    - $X_t \sim norm(0, \sigma^2)$

```r
x = rnorm(1000,0,1)
ts.plot(x, main = "Simulated Gaussian white noise process")
```



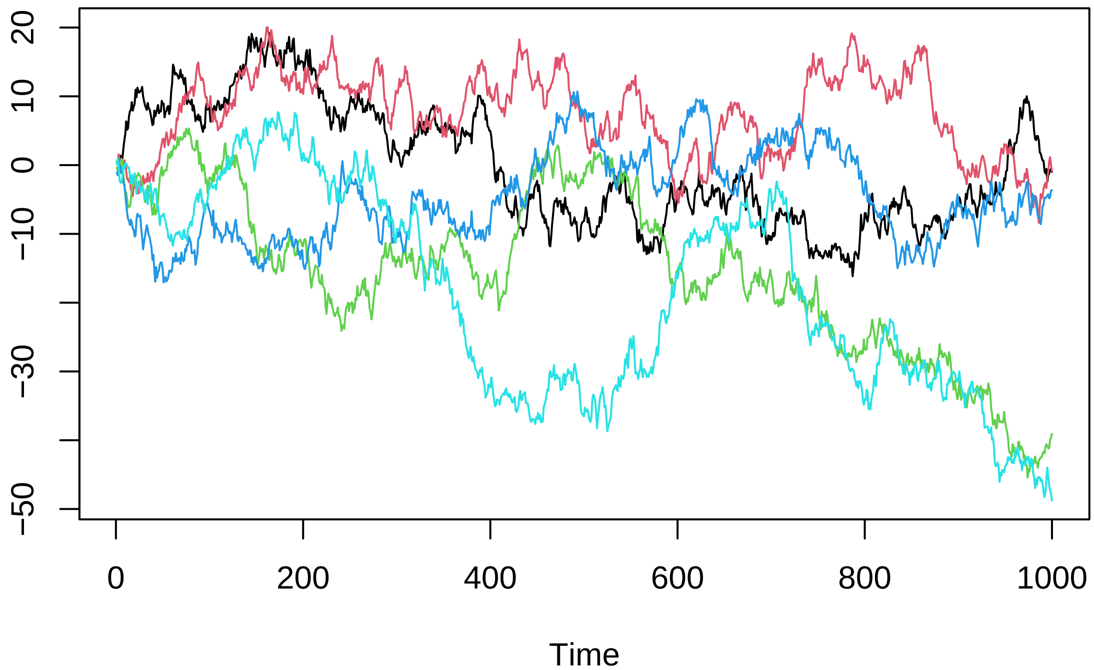
- White noise processes are the simplest stochastic processes.
- Real data does typically not have complete independence between measurements at different time points, so white noise is generally not a good model for real data, but it is a building block for more complicated stochastic processes.

---

3 Example 2: Random walk

- A **random walk** is defined by $X_t = X_{t-1} + W_t$, where W_t is white noise.
- Here are 5 simulations of a random walk:

```
x = matrix(0,1000,5)
for (i in 1:5) x[,i] = cumsum(rnorm(1000,0,1))
ts.plot(x,col=1:5)
```



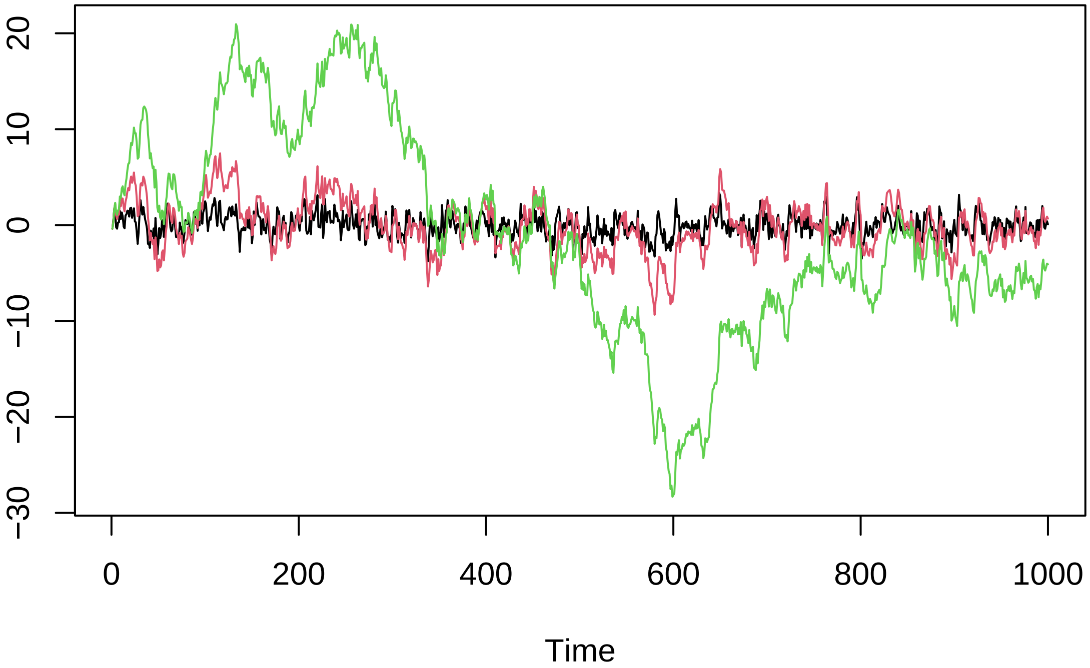
- The random walk may come back to zero after some time, but often it has a tendency to wander of in some random direction.

---

4 Example 3: First order autoregressive process

- A **first order autoregressive process**, AR(1), is defined by $X_t = \alpha X_{t-1} + W_t$, where W_t is white noise and $\alpha \in \mathbb{R}$.
  - Typically $-1 \leq \alpha \leq 1$
  - For $\alpha = 0$ we get white noise
  - For $\alpha = 1$ we get a random walk
- Simulation of 3 AR(1)-processes with different α values:

```r
w = ts(rnorm(1000))
x1 = filter(w,0.5,method="recursive")
x2 = filter(w,0.9,method="recursive")
x3 = filter(w,0.99,method="recursive")
ts.plot(x1,x2,x3,col=1:3)
```



- Next time we will consider autoregressive processes in much more detail and higher order, where they become quite flexible models for data.

5 Mean, autocovariance and stationarity

5.1 Mean function

- The **mean function** of a stochastic process is given by

$$\mu_t = \mathbb{E}(X_t)$$

- A process is called first order stationary if $\mu_t = \mu$.
- **Examples:**
  - The white noise process: $\mu_t = 0$ by definition.
  - The random walk:

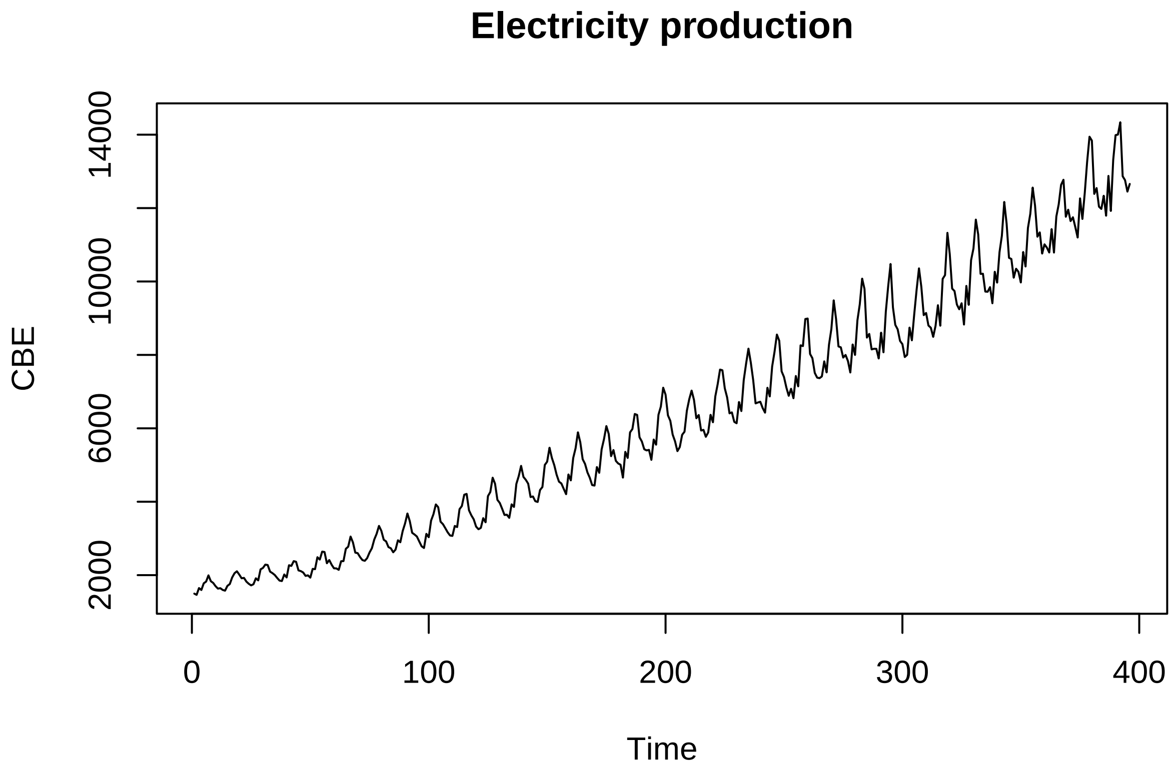
$$\mu_t = \mathbb{E}(X_t) = \mathbb{E}(X_{t-1} + W_t) = \mathbb{E}(X_{t-1}) + \mathbb{E}(W_t) = \mathbb{E}(X_{t-1}) = \mu_{t-1}$$

    So the random walk is first order stationary.
  - Similarly,

$$\mu_t = \mathbb{E}(X_t) = \mathbb{E}(\alpha X_{t-1} + W_t) = \alpha\mathbb{E}(X_{t-1}) + \mathbb{E}(W_t) = \alpha\mathbb{E}(X_{t-1}) = \alpha\mu_{t-1}$$

    The AR(1)-model is first order stationary if $\mu_0 = 0$ or $\alpha = 1$, otherwise not.
  - The electricity production in Australia did not look first order stationary.

```
plot(CBE,main="Electricity production")
```



- The mean function shows the mean behavior of the process, but individual simulations may move far away from this. For example, the random walk has a tendency to move far away from the mean. White noise on the other hand will stay close to the mean.

5.2 Autocovariance/autocorrelation functions

- The **autocovariance** function is given by

$$\gamma(t, t+h) = \mathrm{Cov}(X_t, X_{t+h}) = \mathbb{E}((X_t - \mu_t)(X_{t+h} - \mu_{t+h}))$$

- h is called the **lag**.
- Note that

$$\gamma(t, t) = \mathrm{Var}(X_t) = \sigma_t^2$$

  is the variance at time t.
- The **autocorrelation function (ACF)** is

$$\rho(t, t+h) = \mathrm{Cor}(X_t, X_{t+h}) = \frac{\mathrm{Cov}(X_t, X_{t+h})}{\sigma_t \sigma_{t+h}}$$

- It holds that $\rho(t, t) = 1$, and $\rho(t, t+h)$ is between -1 and 1 for any h.
- The autocorrelation function measures how correlated X_t and X_{t+h} are related:
  - If X_t and X_{t+h} are independent, then $\rho(t, t+h) = 0$
  - If $\rho(t, t+h)$ is close to one, then X_t and X_{t+h} tends to be either high or low at the same time.
  - If $\rho(t, t+h)$ is close to minus one, then when X_t is high X_{t+h} tends to be low and vice versa.

5.3 Stationarity

- We call a stochastic process **second order stationary** if

- the mean is constant, $\mu_t = \mu$
- the variance $\sigma_t^2 = \mathrm{Var}(X_t, X_t)$ is constant.
- the autocorralation function only depends on the lag h:

$$\rho(t, t+h) = \rho(h)$$

- If a process is second order stationary, then also the autocovariance is stationary $\gamma(t, t+h) = \gamma(h)$, i.e. it is a function of only the lag and is easier to work with and plot.
- Intuitively, stationarity means that the process behaves in the same way no matter which time we look at.
- There are other kinds of stationarity, but *in this course, stationarity will always mean second order stationarity.*

5.4 Stationarity and autocorrelation - example

- Consider an AR(1) process $X_t = \alpha X_{t-1} + W_t$. We consider stationarity and autocorrelation for this process.
  - We have already seen that we need $\mu_t = 0$ to have first order stationarity.
- Now consider the variance. Since $X_t = \alpha X_{t-1} + W_t$,

$$\sigma_t^2 = \mathrm{Var}(X_t) = \mathrm{Var}(\alpha X_{t-1} + W_t) = \mathrm{Var}(\alpha X_{t-1}) + \mathrm{Var}(W_t) = \alpha^2 \mathrm{Var}(X_{t-1}) + \mathrm{Var}(W_t) = \alpha^2 \sigma_{t-1}^2 + \sigma^2$$

  - (Here we used that $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$ when X and Y are independent).
- If the variance is constant, then $\sigma_t^2 = \sigma_{t-1}^2$ and

$$\sigma_t^2 = \alpha^2 \sigma_t^2 + \sigma^2$$

  - We see that the variance can only be constant if $-1 < \alpha < 1$. In this case $\sigma_t^2 = \frac{\sigma^2}{1-\alpha^2}$.
  - For $|\alpha| \geq 1$, the variance will increase over time. The process is cannot be stationary (including random walk).
- To find the autocorrelation, first observe

$$X_{t+h} = \alpha X_{t+h-1} + W_{t+h} = \cdots = \alpha^h X_t + \sum_{i=0}^{h-1} \alpha^i W_{t+h-i}$$

- Then we find the autocovariance:

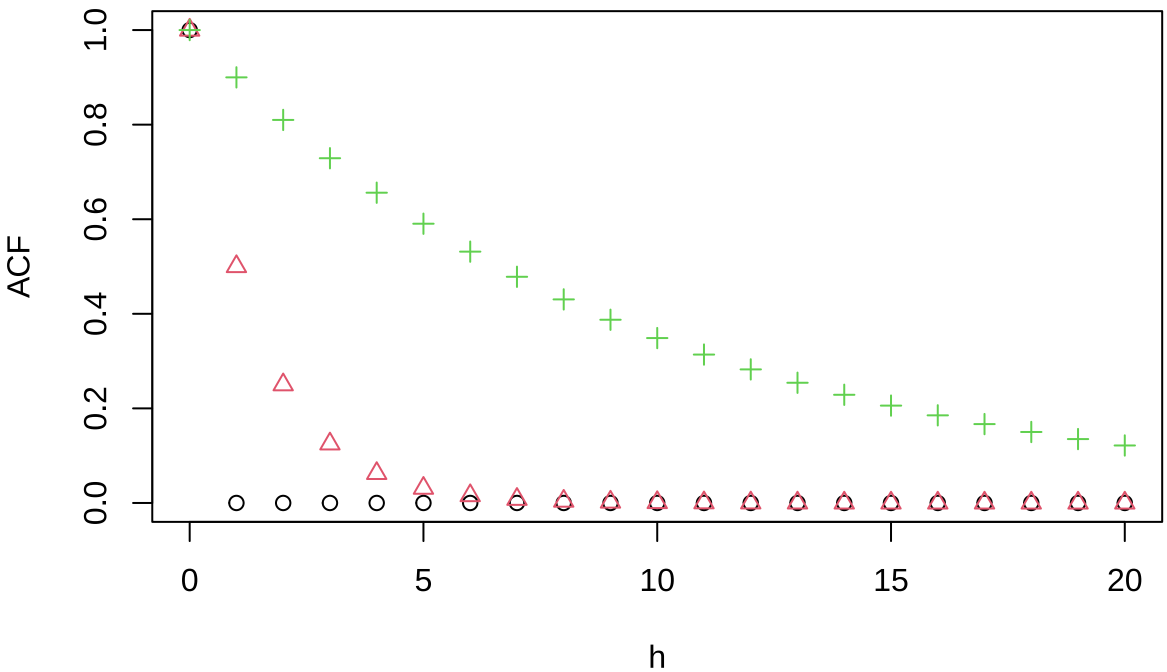
$$\gamma(t, t+h) = \mathrm{Cov}(X_t, X_{t+h}) = \mathrm{Cov}(X_t, \alpha^h X_t + \sum_{i=0}^{h-1} \alpha^i W_{t+h-i}) = \mathrm{Cov}(X_t, \alpha^h X_t) + \mathrm{Cov}(X_t, \sum_{i=0}^{h-1} \alpha^i W_{t+h-i}) = \alpha^h \mathrm{Cov}(X_t, X$$

  - (Here we used the computation rules $\mathrm{Cov}(X, Y+Z) = \mathrm{Cov}(X, Y) + \mathrm{Cov}(X, Z)$ and $\mathrm{Cov}(X, aY) = a\mathrm{Cov}(X, Y)$.)
- If the variance is constant, we can calculate the autocorrelation:

$$\frac{\mathrm{Cov}(X_t, X_{t+h})}{\sigma_t \sigma_{t+h}} = \frac{\alpha^h \sigma^2/(1-\alpha^2)}{\sigma^2/(1-\alpha^2)} = \alpha^h.$$

- So: the AR(1)-model is stationary if $-1 < \alpha < 1$ and $\sigma_t^2 = \sigma^2/(1-\alpha^2)$ - otherwise not.
- The autocorrelation decays exponentially for a stationary AR(1)-model. This is illustrated for 3 different α values:

```
h = 0:20
acf1 = 0^h  # AR(1) with alpha = 0 (or white noise)
acf2 = 0.5^h  # AR(1) with alpha = 0.5
acf3 = 0.9^h  # Ar(1) with alpha = 0.9
plot(matrix(rep(h,3),3),cbind(acf1,acf2,acf3),col=rep(1:3,each=length(h)),
     pch=rep(1:3,each = length(h)),xlab="h",ylab="ACF")
```



6 Estimation

6.1 Estimation

- The mean and autocovariance/autocorrelation functions are theoretical constructions defined for stochastic processes, but what about data? Here we have to estimate them.
- We will assume that the process is stationary.
- The (constant) mean can be estimated the usual way:

$$\hat{\mu} = \bar{x} = \frac{1}{n}\sum_{t=1}^{n} x_t$$

- The autocovariance function can be estimated as follows (remember it only depends on h, not on t in the case of stationarity):

$$\hat{\gamma}(h) = \frac{1}{n}\sum_{t=1}^{n-h}(x_t - \bar{x})(x_{t+h} - \bar{x})$$

- The (constant) variance is estimated as $\hat{\sigma}^2 = \hat{\gamma}(0)$.
- An estimate of the autocorrelation function is obtained as

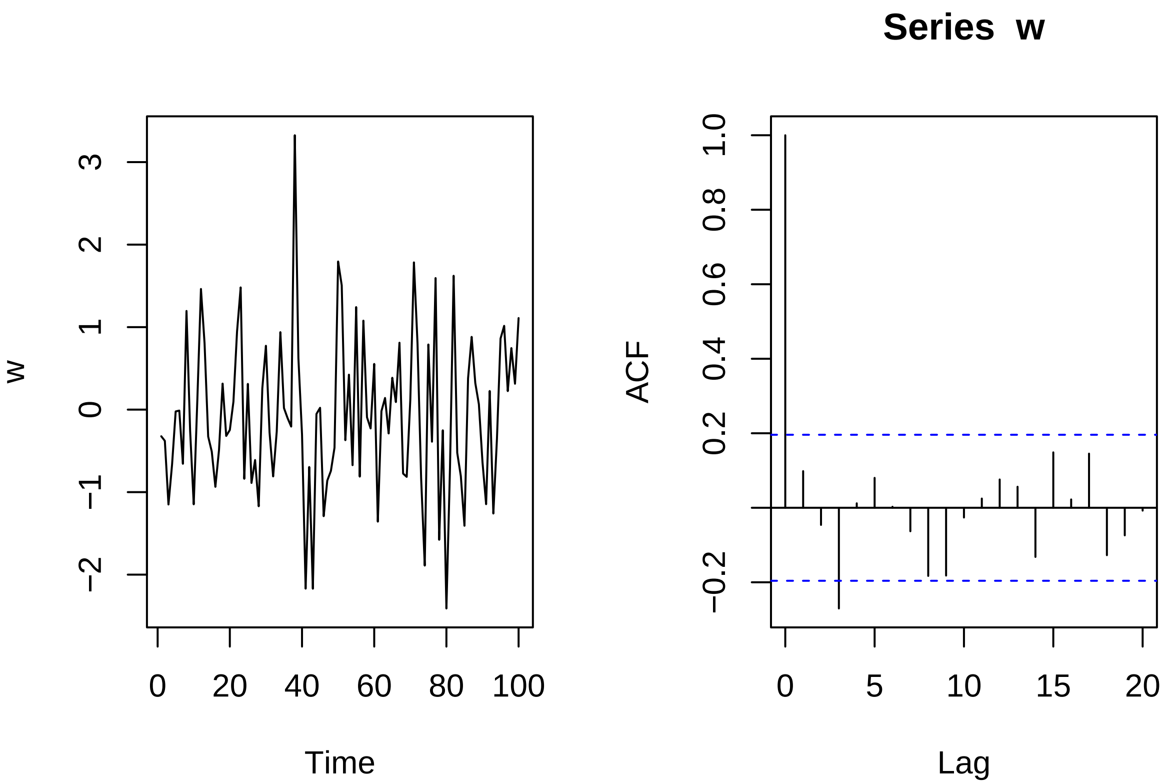
$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}$$

6.2 The correlogram

- A plot of the sample acf as a function of the lag is called a **correlogram**.
- To get an idea of how a correlogram looks, we make simulated data from different models and plot the correlograms below.

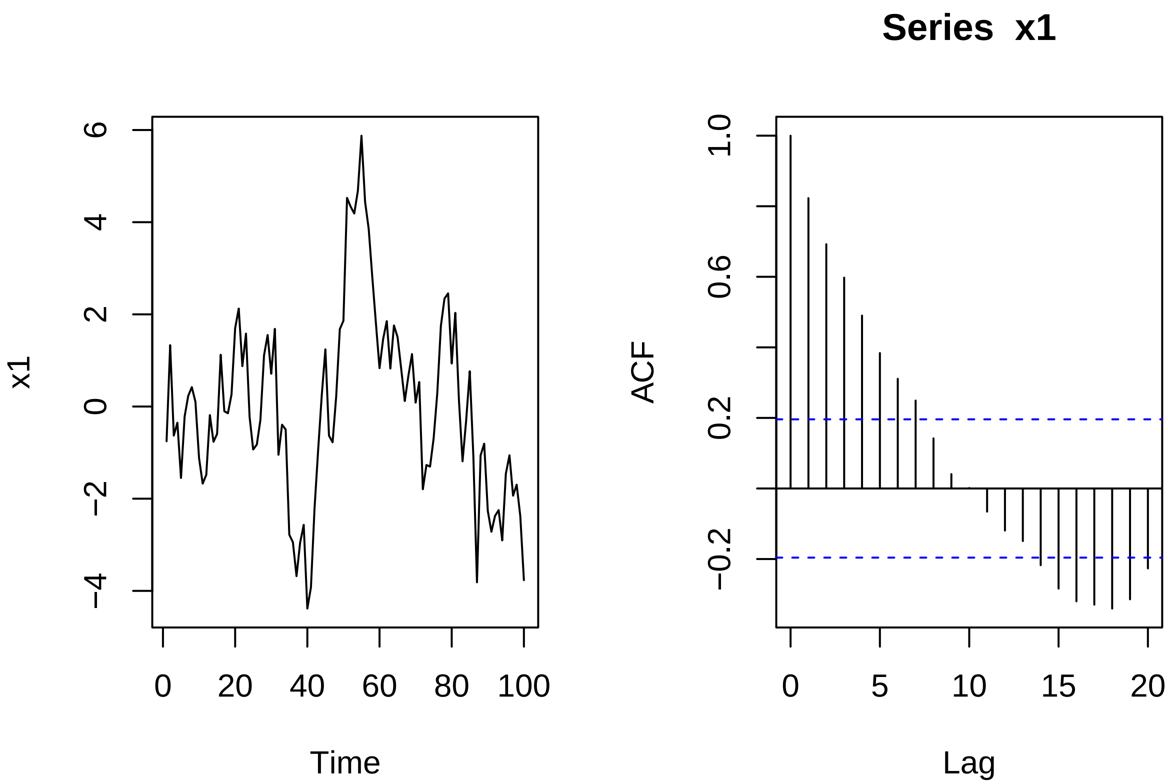
White noise:

```
w = ts(rnorm(100))
par(mfrow=c(1,2))
plot(w)
acf(w)
```



- The correlogram is always 1 at lag 1
- For white noise, the true autocorrelation drops to zero.
- The estimated autocorrelation is never exactly zero - hence we get the small bars.
- The blue lines is a 95% confidence band for a test that the true autocorrelation is zero.
- Remember that there is 5% chance of rejecting a true null hypothesis. Thus, 5% of the bars can be expected to exeed the blue lines.
- AR(1) process with $\alpha = 0.9$:

```
w = ts(rnorm(100))
x1 = filter(w,0.9,method="recursive")
par(mfrow=c(1,2))
plot(x1)
acf(x1)
```



- The true acf decays exponentially.

7 Non-stationary data

7.1 Check for stationarity

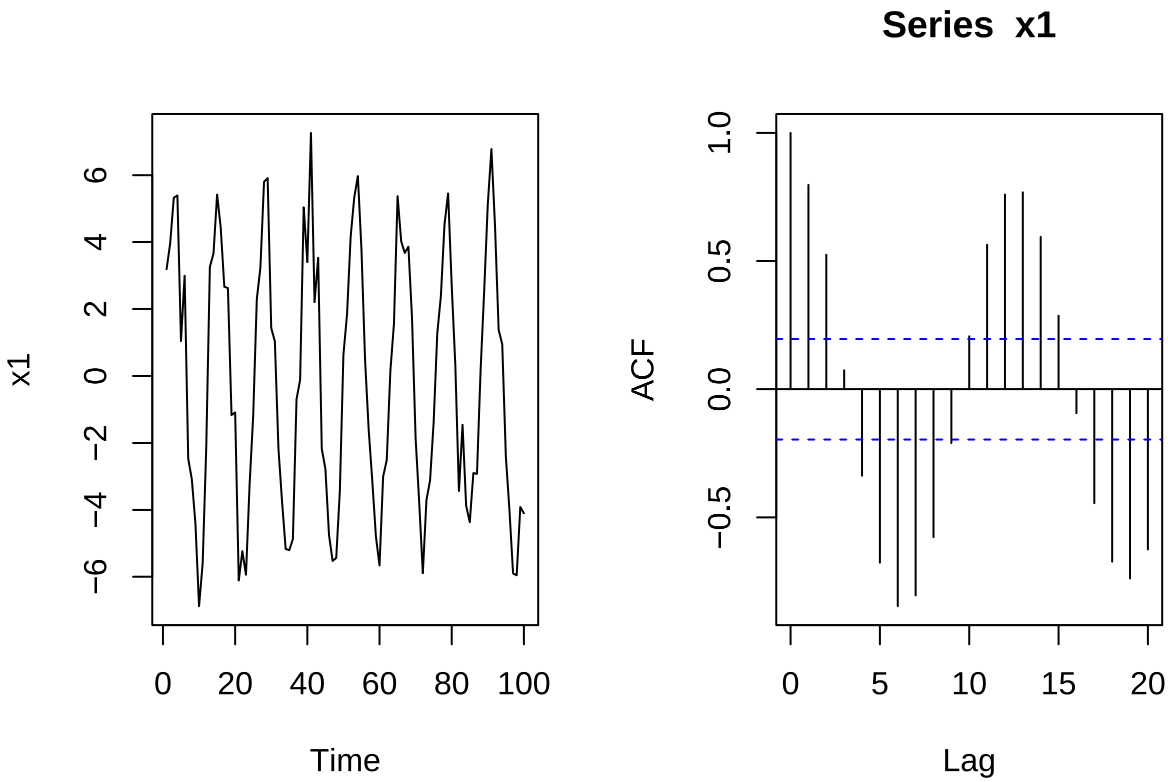
- We will primarily look at stationary processes the next time, but these will not always be good models for data.
- First we need to check whether the assumption of stationarity is okay.
  - One check is visual inspection of a plot of x_t vs t to see whether there is any indication of non-stationarity.
  - Another visual check is a plot of the correlogram. If this tends very slowly to zero, this indicates non-stationarity.
- Note: even though $\rho(h)$ is only well-defined for stationary models, we can plug any data (stationary or not) into the estimation formula. The estimate may help detecting deviations from stationarity.

7.2 Correlograms for non-stationary data

- Sine curve with added white noise:

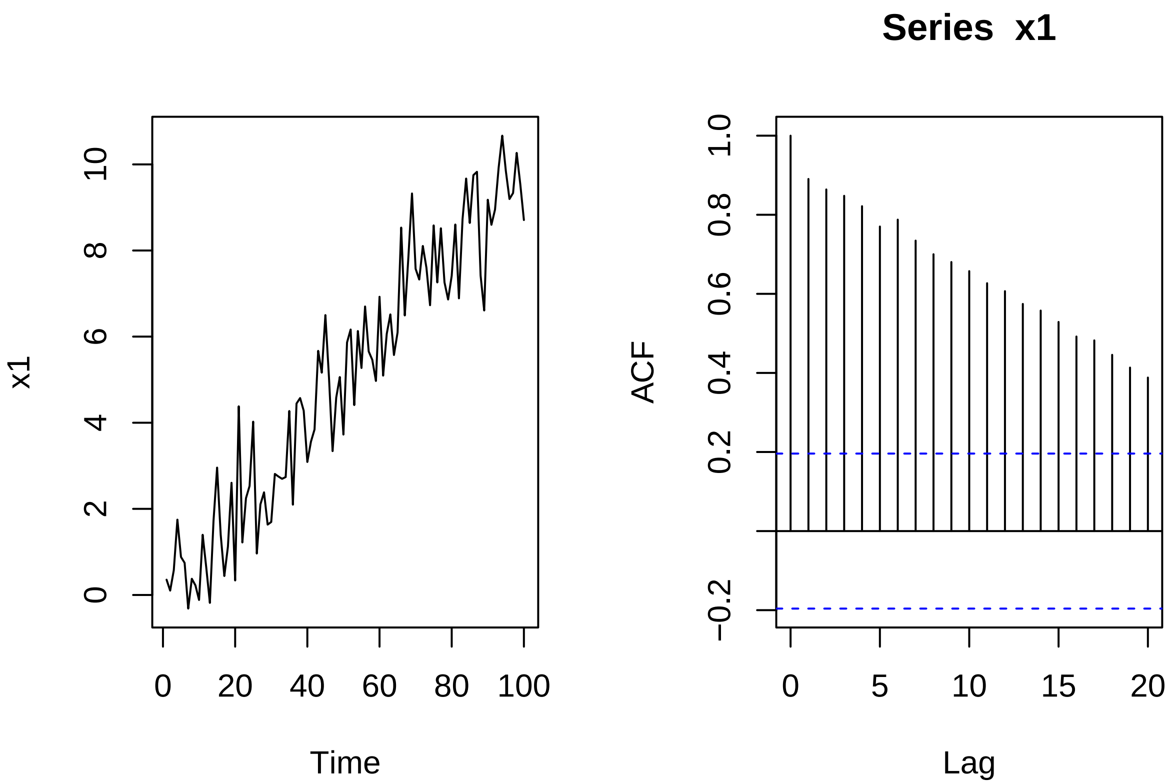
```
w = ts(rnorm(100))
x1 = 5*sin(0.5*(1:100)) + w
par(mfrow=c(1,2))
```

```
plot(x1)
acf(x1)
```



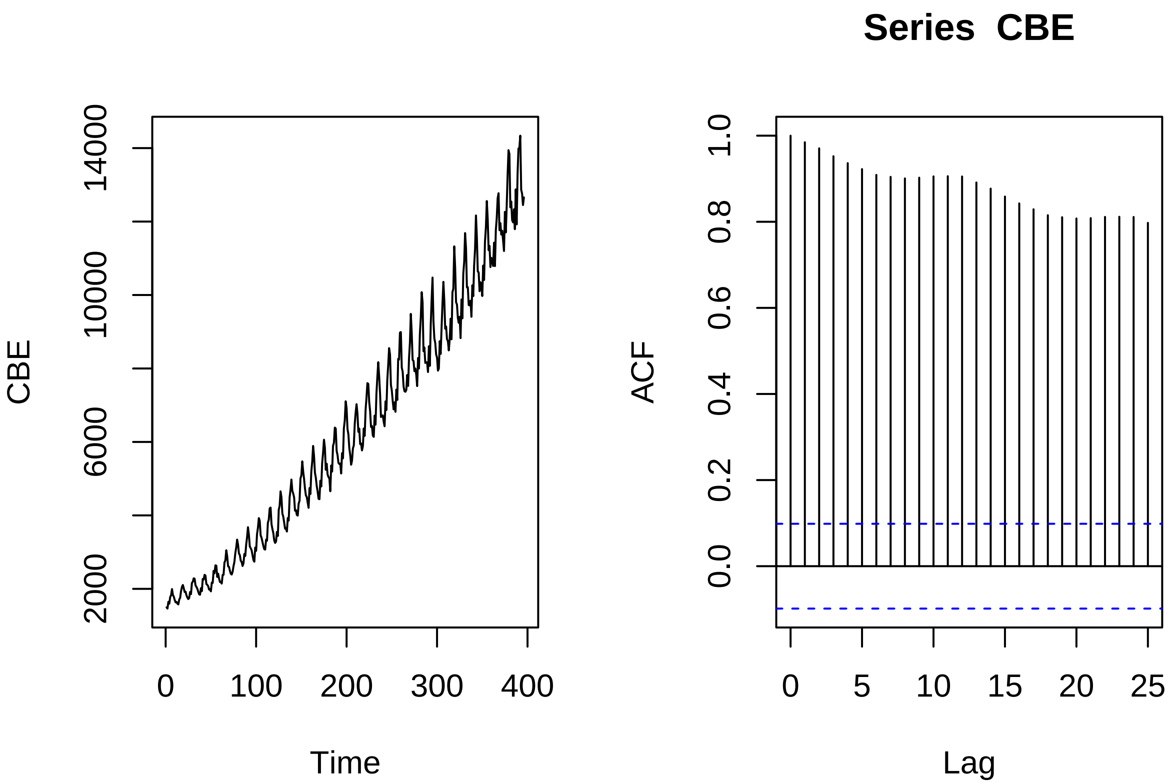
- The periodic mean of the process results in a periodic behavior of the correlogram.
- A periodic behavior in the correlogram suggests seasonal behavior in the process.
- Straight line with added white noise:

```
w = ts(rnorm(100))
x1 = 0.1*(1:100) + w
par(mfrow=c(1,2))
plot(x1)
acf(x1)
```



- The linear trend results in a slowly decaying, almost linear correlogram.
- Such a correlogram suggests a trend in the data.
- Data example: Electricity production.

```
par(mfrow=c(1,2))
plot(CBE)
acf(CBE)
```



- There seems to be an increasing trend in the data.
- There is a periodic behavior around the increasing trend.
- It is reasonable to believe that the period is 12 months.
- We have the model

$$X_t = m_t + s_t + Z_t$$

  where

  - m_t is the (deterministic) trend
  - s_t is a (deterministic) seasonal term ($s_t = s_{t+12}$)
  - Z_t is a random (hopefully) stationary part

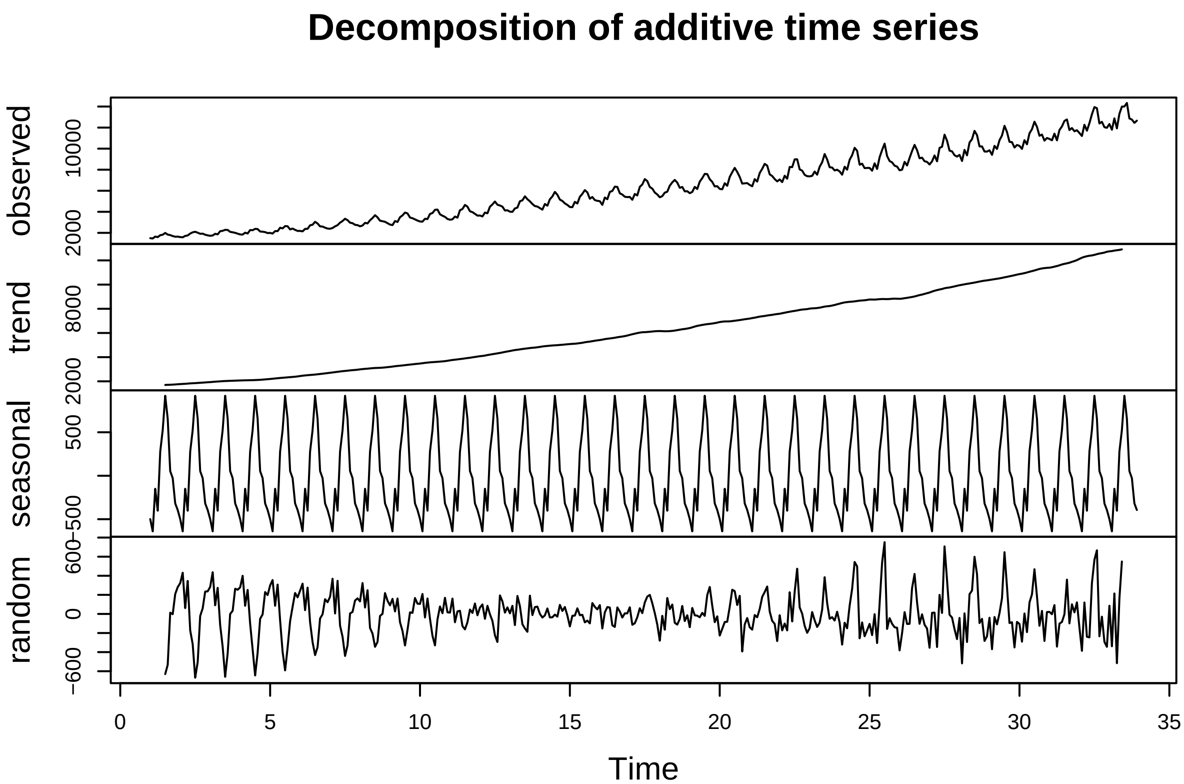
7.3 Detrending data

- The trend m_t in the data can be estimated by a **moving average**.
- In the case of monthly variation,

$$\hat{m}_t = \frac{\frac{1}{2}x_{t-6} + x_{t-5} + \cdots + x_t + \cdots + x_{t+5} + \frac{1}{2}x_{t+6}}{12}$$

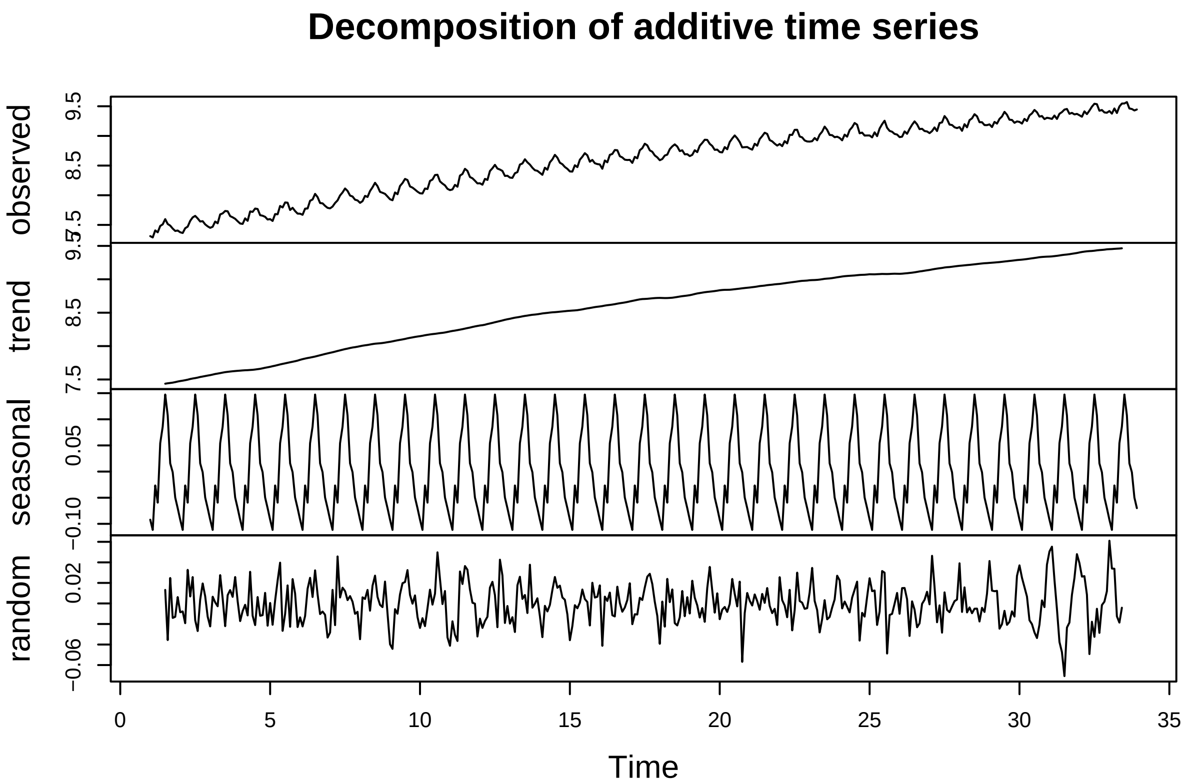
- We remove the trend by considering $x_t - \hat{m}_t$.
- Next we find the seasonal term s_t by averaging $x_t - \hat{m}_t$ over all measurements in the given month.
  - E.g., the value of s_t for January is given by averaging all values from January.
- We are left with the random part $\hat{z}_t = x_t - \hat{m}_t - \hat{s}_t$.
- For the Australian electricity data:

```r
CBE <- ts(CBE,frequency=12)
plot(decompose(CBE))
```



- The random term does not look stationary. The solution is to log-transform the data - see Ch. 1.5.5 in the book.

```r
logCBE <- ts(log(CBEdata[,3]),frequency=12)
plot(decompose(logCBE))
```




```
random<-decompose(logCBE)$random[7:382]
acf(random, main="Random part of CBE data")
```