Stochastic processes I

The ASTA team

Contents

1 Data examples

- Before we define a stochastic process, we start by considering some examples of data.
- Here we are considering data which is a collection of the same variable measured at different points in time, i.e. x_t indexed by time t in some discrete set, often $t = 1, \ldots, n$. We will always assume the data is observed at equidistant points in time (i.e. same time difference between consecutive observations).
- An example is the monthly number of international airline passengers 1949-1960:

```
AP <- AirPassengers
plot(AP, ylab = "Bookings (1000s)")
```
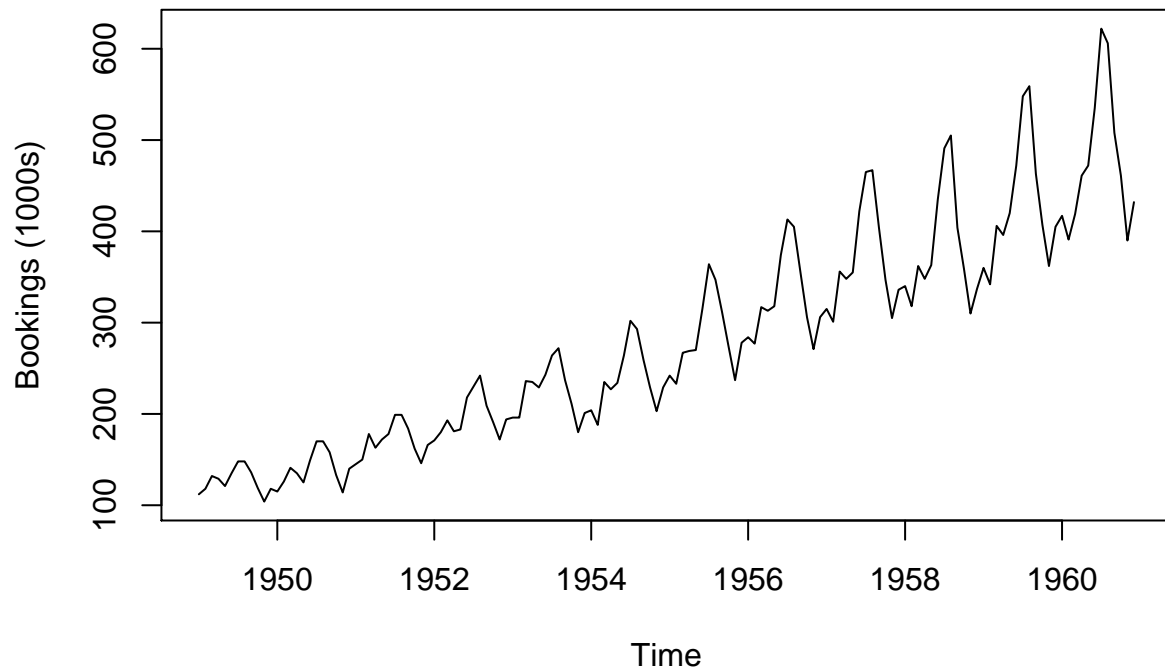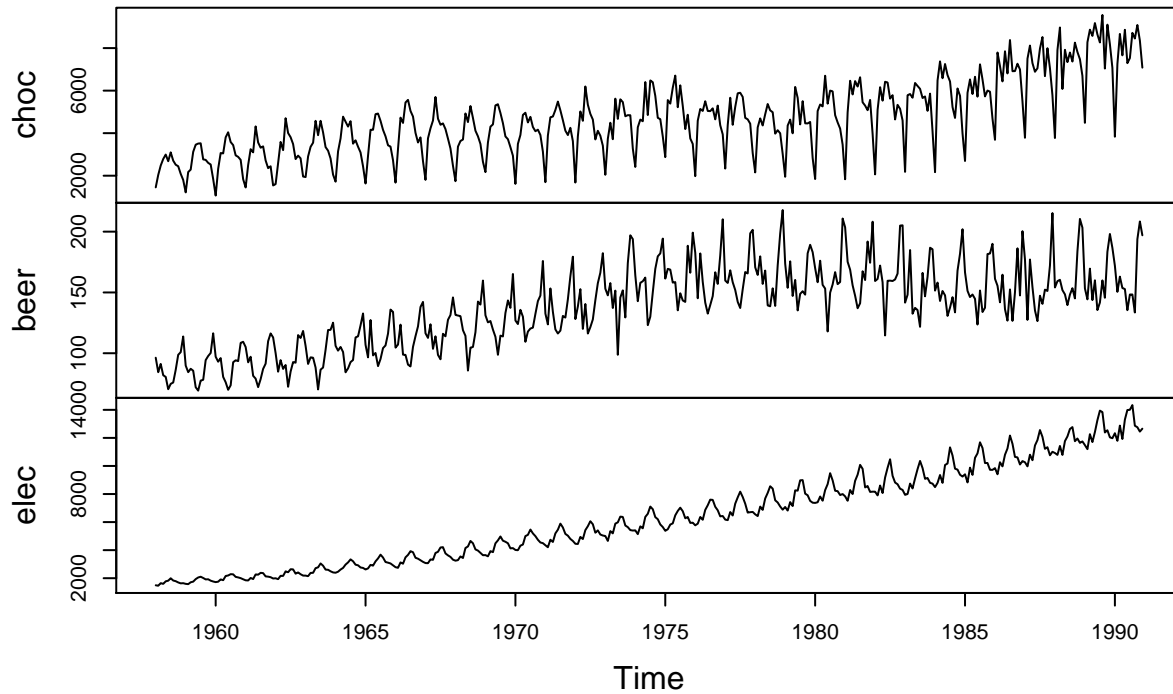
- Another example is monthly time series from Jan. 1958 to Dec. 1990 of supply of three goods in Australia:

  - Electricity (Giga Watt hours)
  - Beer (Mega liters)
  - Chocolate (tonnes)

```
CBEdata <- read.table("https://asta.math.aau.dk/eng/static/datasets?file=cbe.dat", header = TRUE)
CBE <- ts(CBEdata, start = 1958, freq = 12)
plot(CBE)
```

CBE

- There are many examples of data evolving over time and measured at discrete times, and we would like to make realistic models to analyse and make predictions for such datasets.

2 Stochastic processes

- A stochastic process is a number of stochastic variables X_t indexed by time t, for example $t = 1, \ldots, n$ or $t \in \mathbb{Z}$.
- Typically there will be dependence between X_t at different time points.
- We will often define a stochastic process by defining the distribution of X_t conditionally on X_s for $s < t$.

3 Example 1: White noise

- White noise is the simplest example of a stochastic process: Here X_t are independent and identically distributed random variables with mean 0 and variance σ^2. It is called Gaussian white noise, if X_t is Gaussian.
- Due to independence everywhere, white noise is typically not a good model for real data, but it is a building block for more complicated stochastic processes.
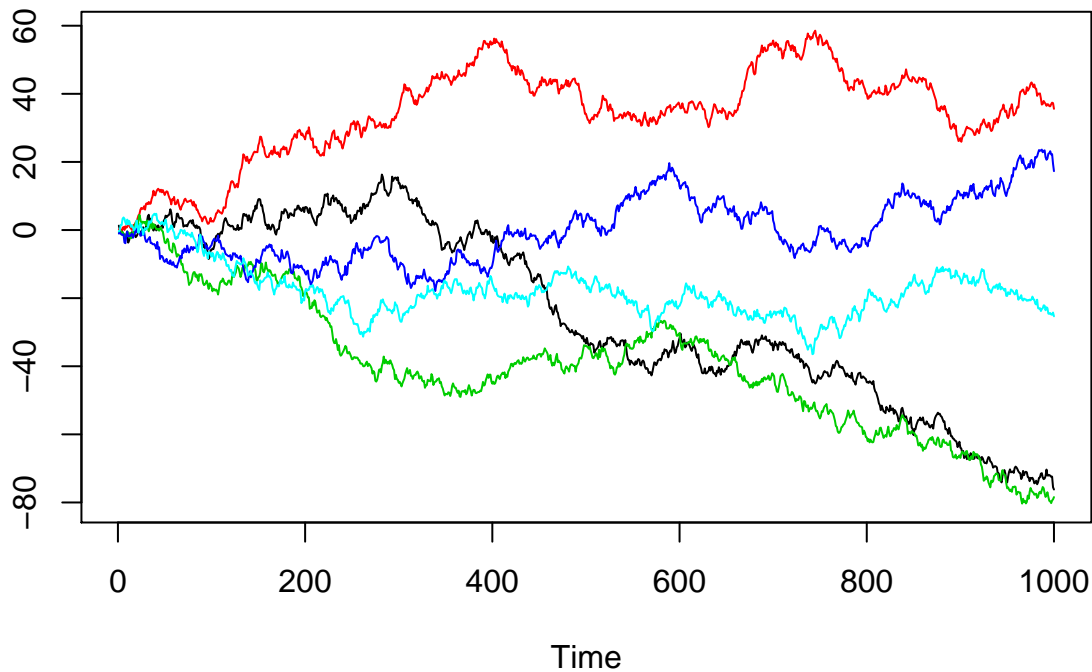
```
x = rnorm(1000,0,1)
ts.plot(x)
```
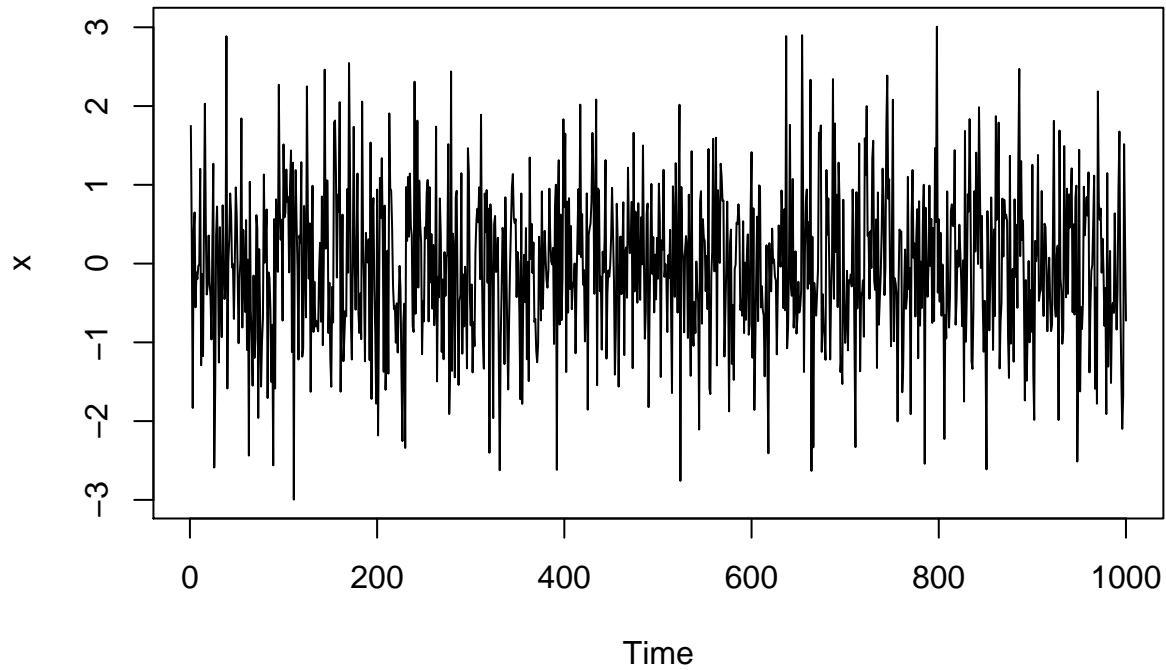
4 Example 2: Random walk

- A random walk is defined by $x_t = x_{t-1} + w_t$, where w_t is white noise.
- The random walk may come back to zero after some time, but often it has a tendency to wander of in some random direction.
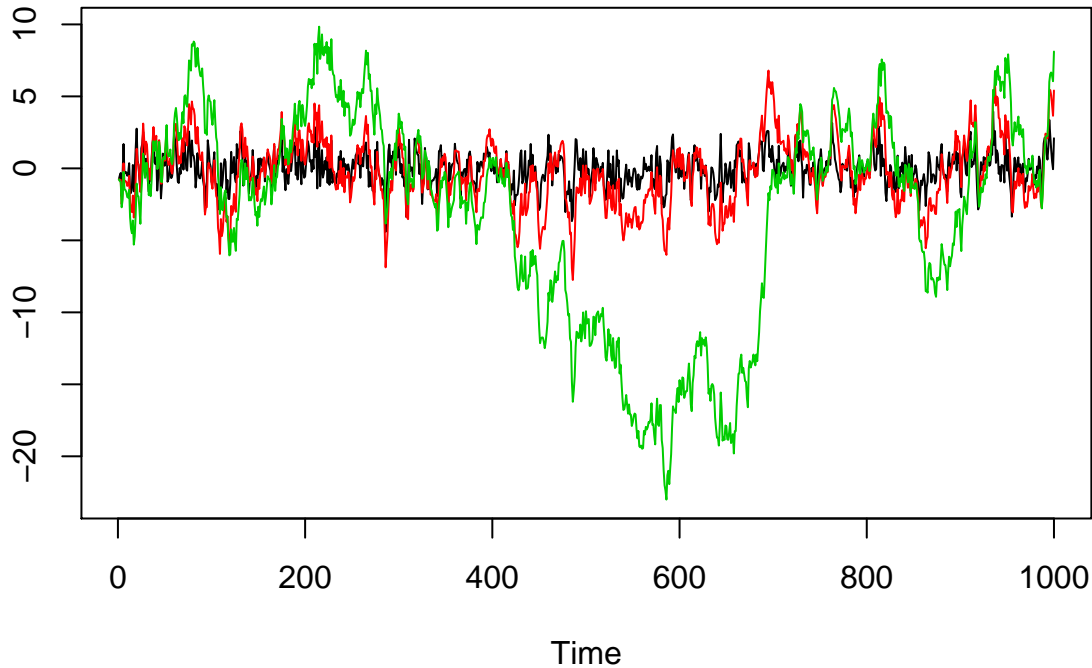- Examples:

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

Time

5 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 $\alpha \in \mathbb{R}$.
- For $\alpha = 0$ we get white noise, and for $\alpha = 1$ we get a random walk.
- Next time we will consider autoregressive processes in much more detail and higher order, where they become quite flexible models for data.
- Examples:

```
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)
```

6   Mean function

- The mean function of a stochastic process is given by

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

- All three examples have a constant mean of $\mu_t = 0$. For example, the random walk:

$$\mathbb{E}(x_t) = \mathbb{E}(x_{t-1} + w_t) = \mathbb{E}(x_{t-1}) + \mathbb{E}(w_t) = \mathbb{E}(x_{t-1})$$

$$\Rightarrow \quad \mathbb{E}(x_t) = \mathbb{E}(x_{t-1}) = \cdots = \mathbb{E}(x_0) = \mathbb{E}(0) = 0$$

- 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.

7   Autocovariance/autocorrelation functions

- The autocovariance function is given by

$$\gamma(t, t + h) = \text{Cov}(x_t, x_{t+h}) = \mathbb{E}((x_t - \mu_t)(x_{t+h} - \mu_{t+h}))$$

- Note that $\gamma(t, t) = \sigma_t^2$ is the variance at time t.
- The autocorrelation function (ACF) is a normalised version of the autocovariance function

$$\rho(t, t + h) = \frac{\text{Cov}(x_t, x_{t+h})}{\sigma_t \sigma_{t+h}}$$

- It holds that $\rho(t, t) = 1$, and $\rho(t, t + h)$ is in the interval [-1,1] for any h.
- The autocorrelation function shows how much 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.

8 Stationarity

- We call a stochastic process second order stationary if

- the mean is constant, $\mu_t = \mu$, and
- the autocovariance function only depends on the time difference, $\gamma(t, t+h) = \gamma(h)$.
- In this case the variance $\sigma_t^2 = \gamma(t, t) = \gamma(0)$ is also constant.
- If a process is second order stationary, then also $\rho(t, t+h) = \rho(h)$, i.e. it is a function of only h and is easier to work with and plot.
- Intuitively stationarity means that the process behaves in the same way nomatter which times we look at.
- There are other kinds of stationarity, but in this course by stationarity we always mean second order stationarity.

9 Stationarity and autocorrelation - example

- The autoregressive process is stationary if $\alpha \in (-1, 1)$. We can calculate the autocovariance and 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 calculate the autocovariance:

$$\gamma(h) = \mathbb{E}((x_t - \mu_t)(x_{t+h} - \mu_{t+h})) = \mathbb{E}(x_t x_{t+h}) = \mathbb{E}(x_t(\alpha^h x_t + \sum_{i=0}^{h-1} \alpha^i w_{t+h-i})) = \mathbb{E}(x_t \alpha^h x_t) + \mathbb{E}(x_t \sum_{i=0}^{h-1} \alpha^i w_{t+h-i}) = \alpha^h \mathbb{E}(x_t^2$$

- We need to calculate the variance of x_t, where we use $\mathrm{Var}(x_t) = \mathrm{Var}(x_{t-1})$ by stationarity:

$$x_t = \alpha x_{t-1} + w_t \Rightarrow \mathrm{Var}(x_t) = \mathrm{Var}(\alpha x_{t-1}) + \mathrm{Var}(w_t)$$

$$\Rightarrow \mathrm{Var}(x_t) = \alpha^2 \mathrm{Var}(x_t) + \sigma^2 \Rightarrow \mathrm{Var}(x_t) = \frac{\sigma^2}{1 - \alpha^2}$$

- Finally the autocorrelation:
$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \frac{\alpha^h \sigma^2/(1 - \alpha^2)}{\sigma^2/(1 - \alpha^2)} = \alpha^h$$

- Note that this is only for $h \geq 0$. For arbitrary h we have $\rho(h) = \alpha^{|h|}$ since $\rho(-h) = \rho(h)$.
- White noise is a special case of an autoregressive process with $\alpha = 0$.

$$\rho(h) = 0^{|h|} = \begin{cases} 1, & h = 0 \\ 0, & h \neq 0 \end{cases}$$

- Random walk is not stationary.
- The ACF for a stationary AR(1):

```
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")
```
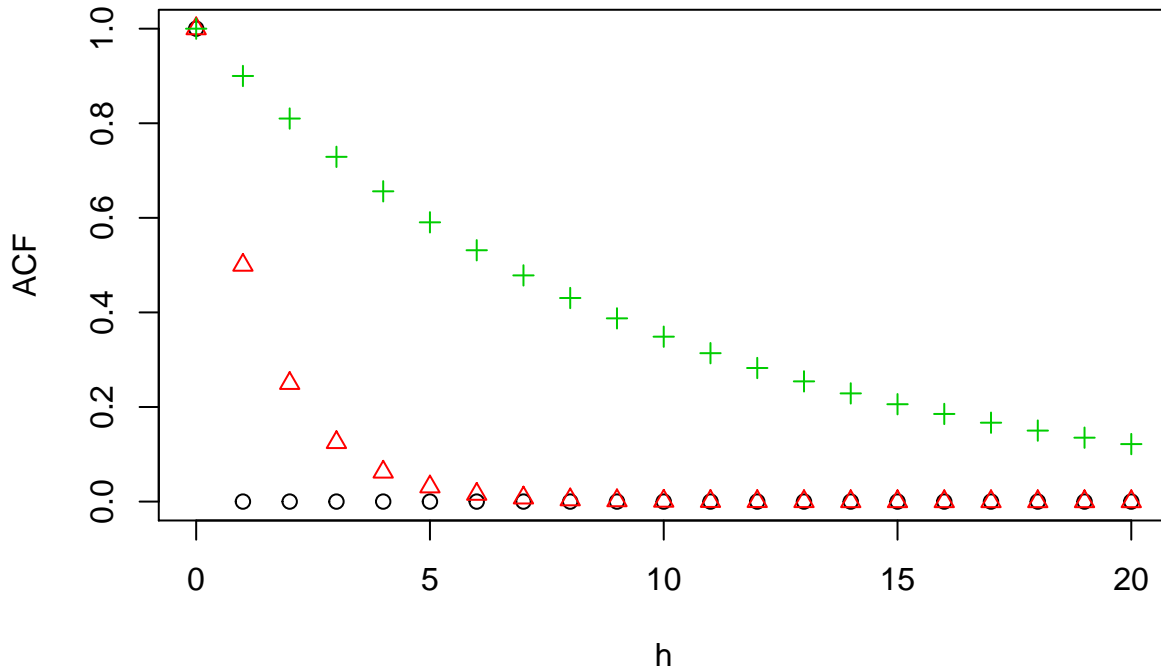
10 Estimation

- 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 focus on the stationary case, i.e. the case where the data looks stationary.
- The (constant) mean can be estimated the usual way:

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

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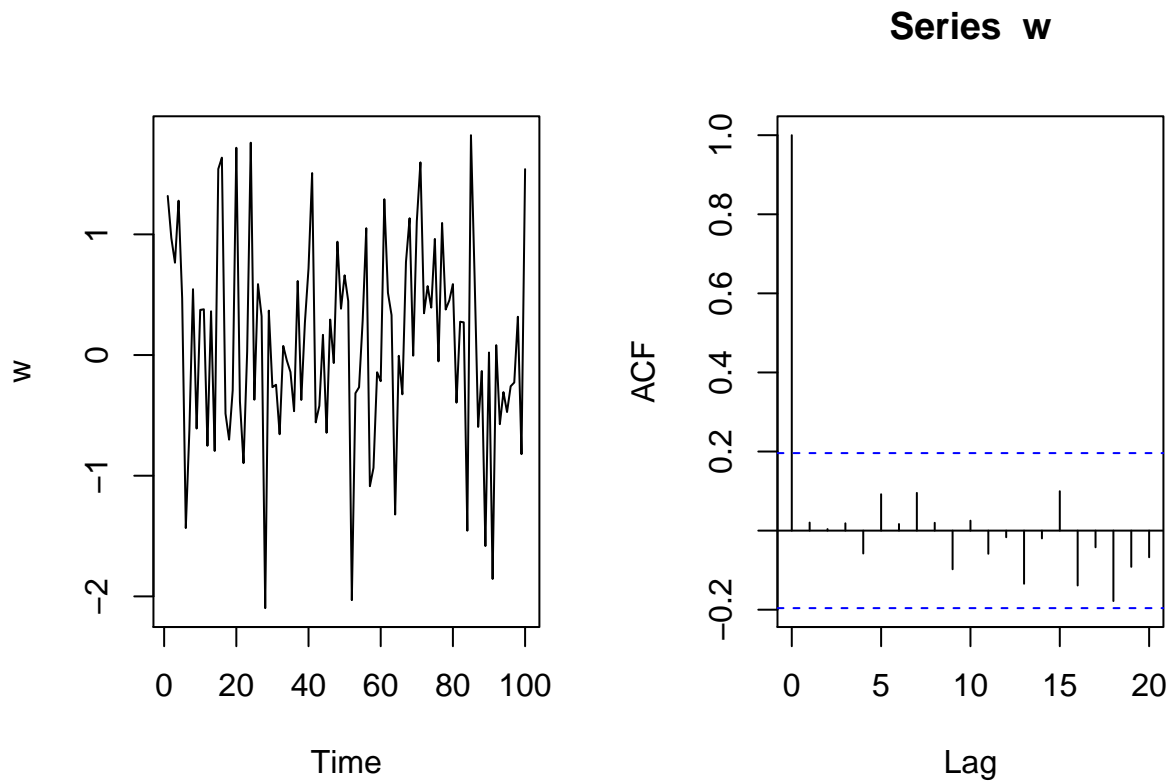- An estimate of the autocorrelation function is obtained as

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

This is known as the correlogram, and has many practical uses.
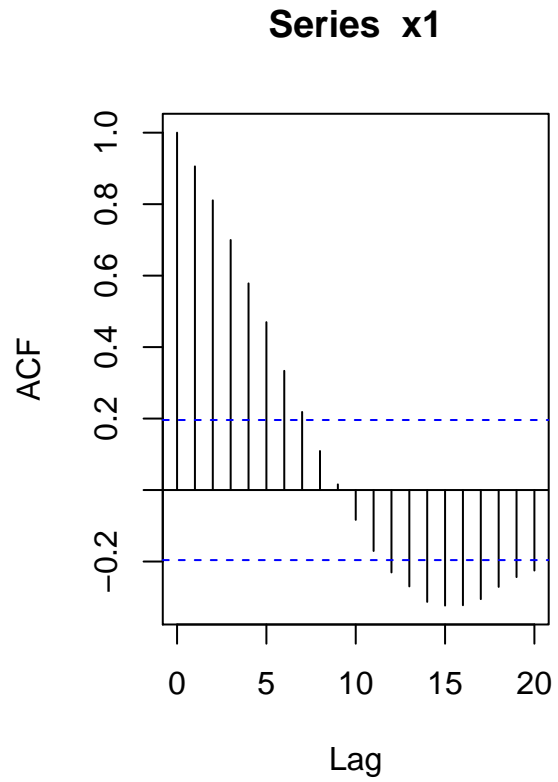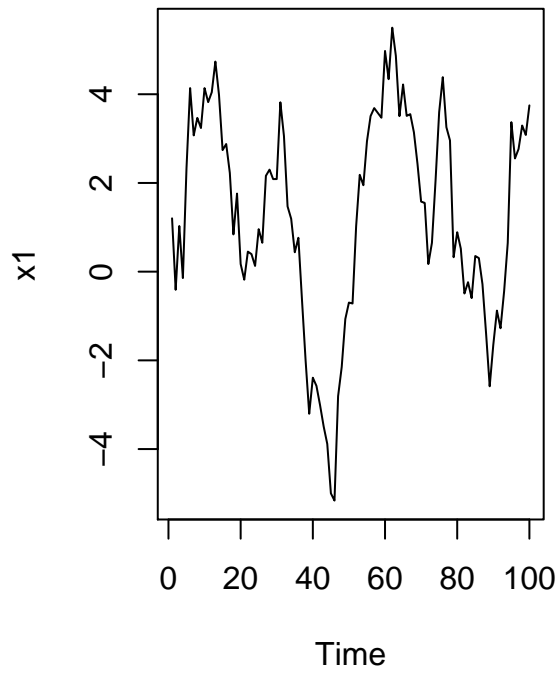
11 Examples of correlograms

- To get an idea of how a correlogram looks, we make simulated data, and use the estimation formulas.
- White noise:

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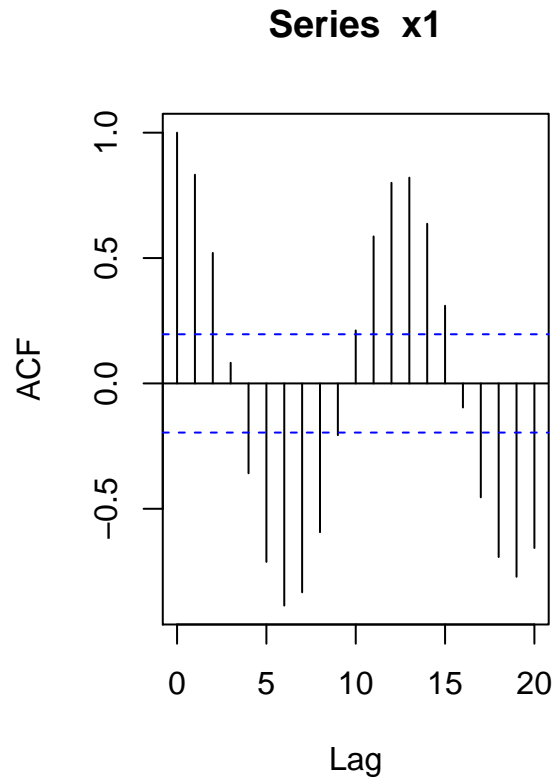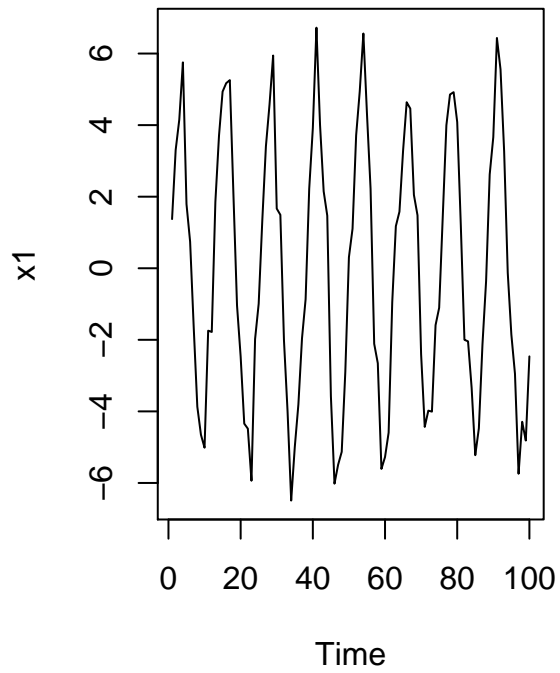
Series w

- 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)
```

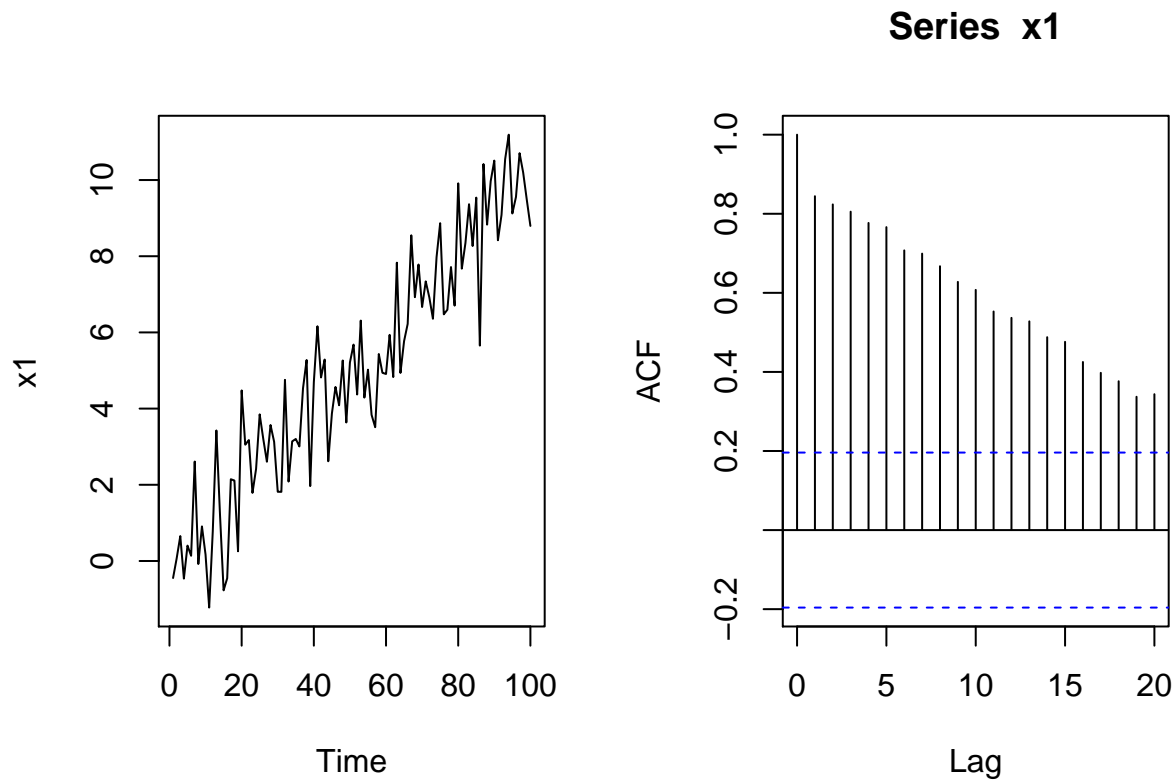**Series x1**

- Sine curve with added white noise:

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

Series x1

- Straight line with added white noise:

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

Series x1

- Note that even though $\rho(h)$ is only well-defined for stationary models, we can plug any data (stationary or otherwise) into the estimation formula. The estimate tells a lot about the data.
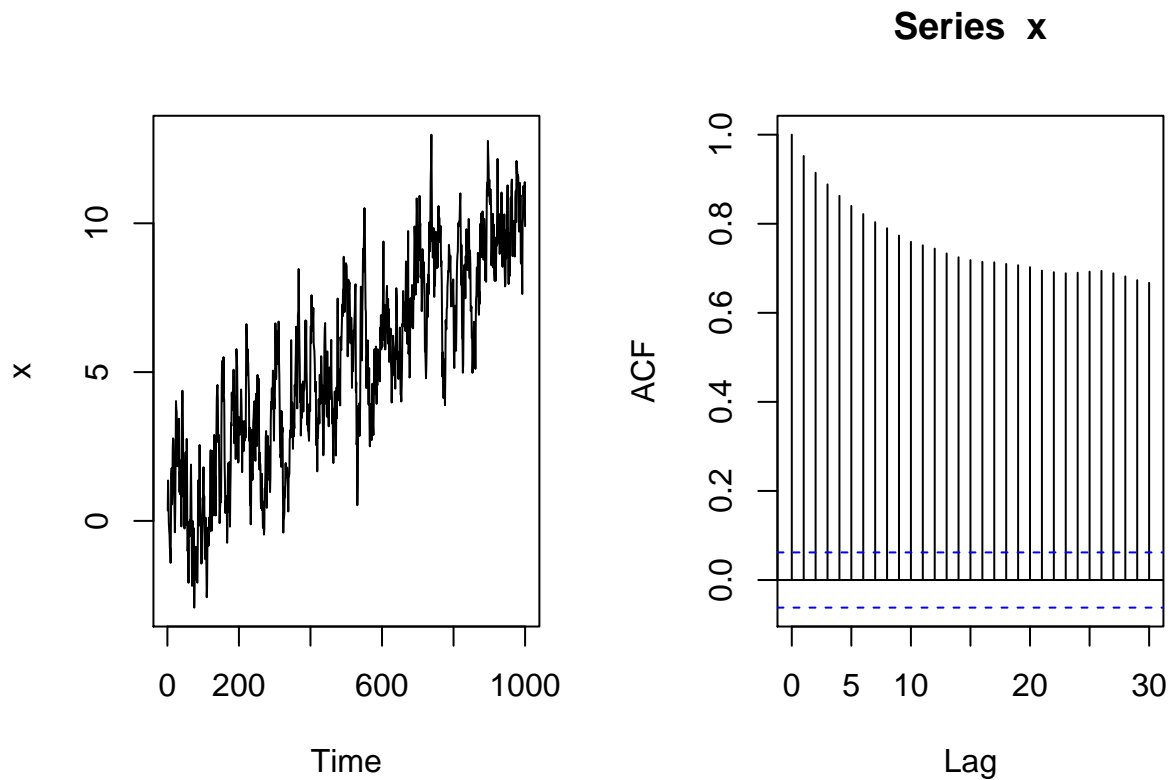
12   Stationary and non-stationary data

- We will primarily look at stationary processes the next time, but these will not be good models for data, if the data does not look stationary. 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.
- Simulated data - stationary:

```
w = ts(rnorm(1000))
x = filter(w,0.8,method="recursive")
par(mfrow=c(1,2)); plot(x); acf(x)
```

Series x
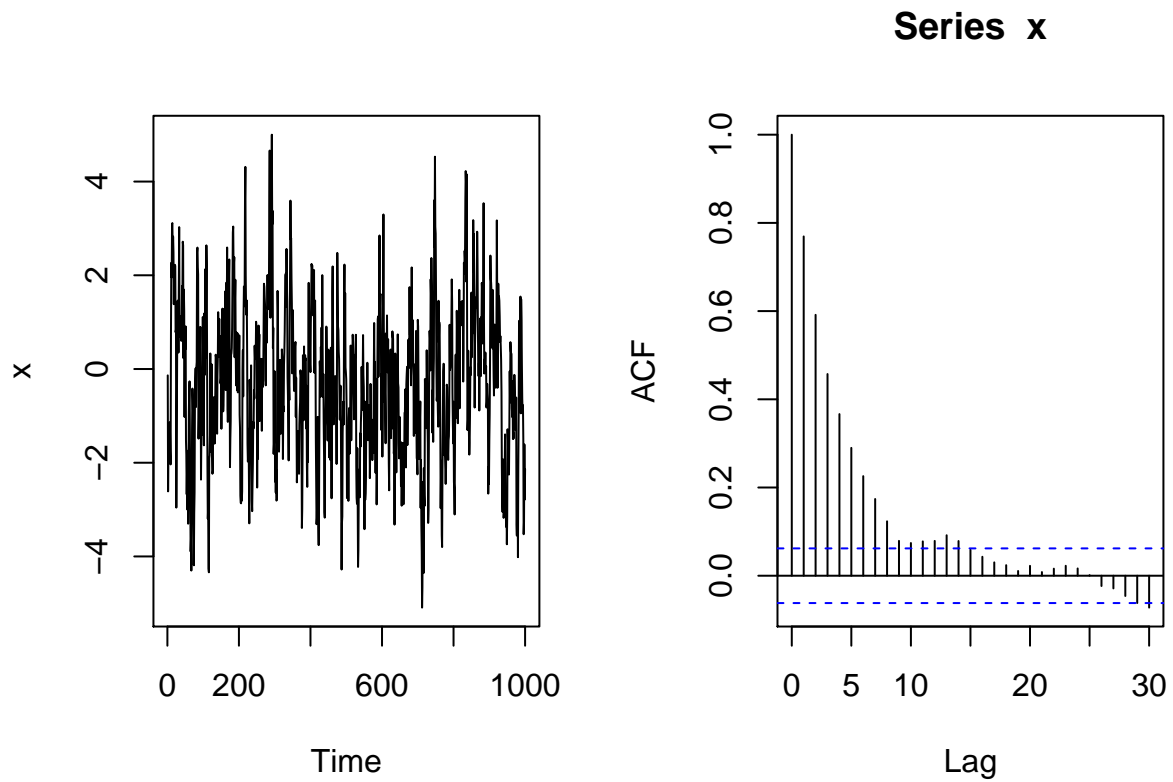
- Non-stationary simulated example (there is an increasing tendency):

```r
w = ts(rnorm(1000))
x = filter(w,0.8,method="recursive") + (1:1000)/100
par(mfrow=c(1,2)); plot(x); acf(x)
```

Series x

13 Detrending data

- If the data does not seem stationary, we can try find a stationary data hidden in it by transforming the data. There are many methods for this, and we will only have a brief look at one possibility based on the linear regression models you have seen earlier.
- Consider the following simulated data:

```
w = ts(rnorm(1000))
x = filter(w,0.9,method="recursive") + (1:1000)/100
par(mfrow=c(1,2)); plot(x); acf(x)
```

Series x
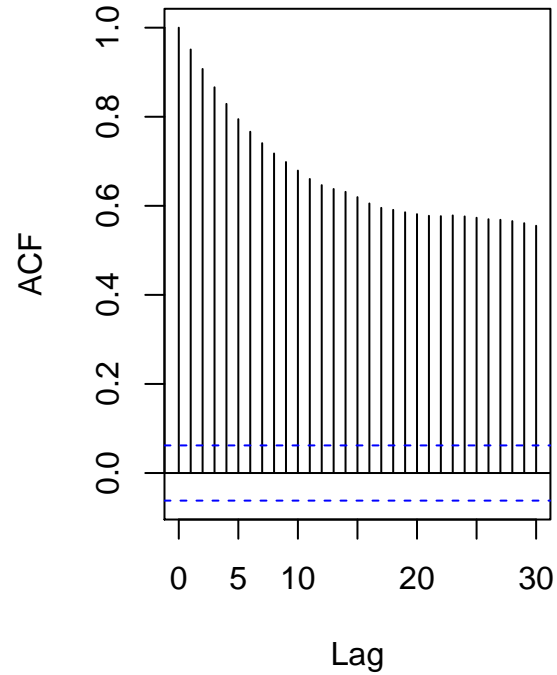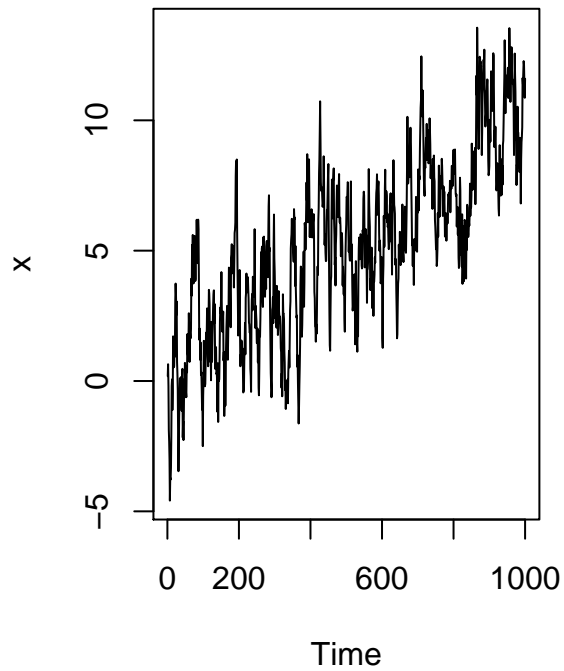
- Ignoring that this is simulated data, it would seem that the data is located around a straight line, so one possibility is to fit a straight line, and the consider the residuals as our new data. We then do all the model fitting to this detrended data, and work with this.

```
lin = lm(x~I(1:1000))
detx = resid(lm(x~I(1:1000)))
par(mfrow=c(1,3)); ts.plot(x); abline(lin,col=2); ts.plot(detx); acf(detx)
```