Lecture 12 – basic properties of random process

Wengang Mao (Marine Technology)
Department of Mechanics and Maritime Sciences,
Chalmers University of Technology,
Goteborg, Sweden

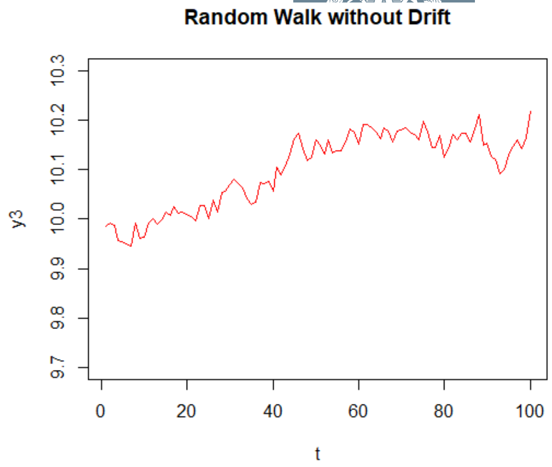
CHALMERS
UNIVERSITY OF TECHNOLOGY

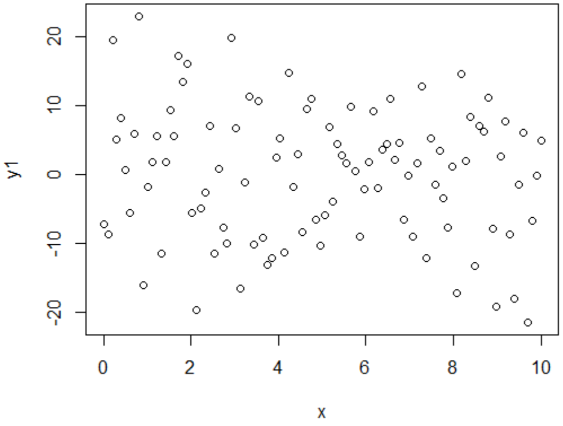
Outline

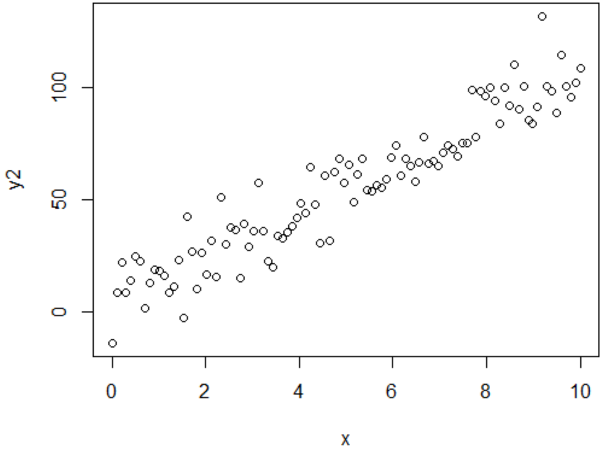
- **Different random data types**
- **Properties to describe random process**
  - Mean, variance
  - Covariance and correlation function
  - Stationary random process
  - Weakly stationary random process
- **Typical examples of random process**
  - White noise
  - Random walk
  - Stationary Gaussian process

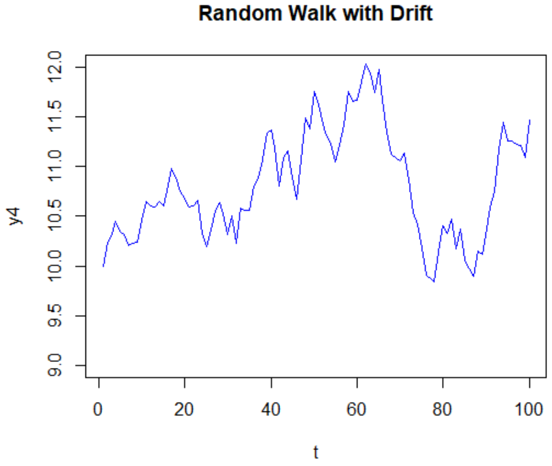
Random data and their models

- A random variable $y_1 \sim N(0, 10)$ that independent of X
- A random variable $y_2 = 5 + 10x + \varepsilon$, where $\varepsilon \sim N(0, 10)$
- A random walk process $y_{3/4}(t) = y_{3/4}(t-1) + \alpha + \varepsilon$, where $\varepsilon \sim N(0, \sigma)$
  - α=0 without drift; $\alpha \neq 0$ with drift

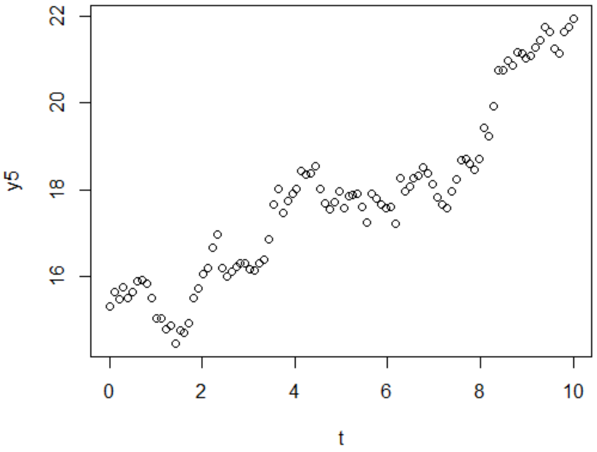




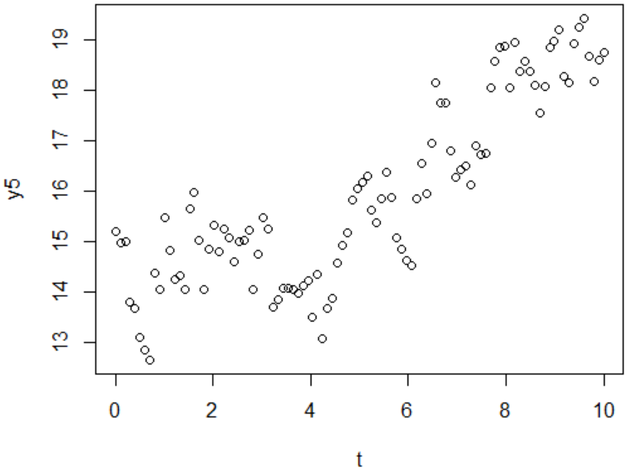




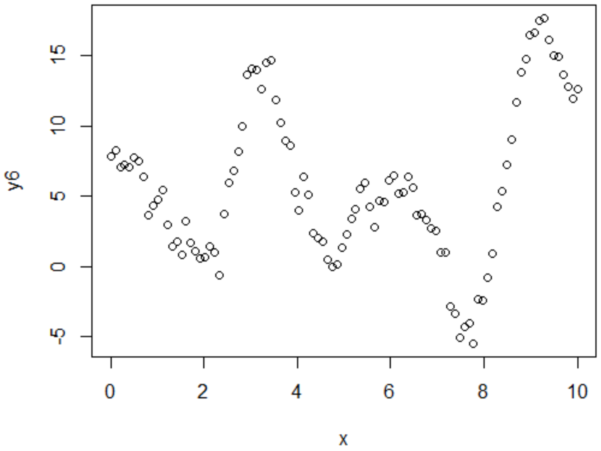
More types of random processes



Random process with trend, i.e.,
$y_5(t) = y_5(t-1) + f(t) + \varepsilon$



Random process with additive model trend,
i.e., $y_5(t) = y_5(t-1) + f(t) + f(x) + \varepsilon$



Random process with period trend,
i.e., $y_6(t) = y_6(t-1) + cos(t) + \varepsilon$

Outline

- Different random data types
- **Properties to describe random process**
  - Mean, variance
  - Covariance and correlation function
  - Stationary random process
  - Weakly stationary random process
- **Typical examples of random process**
  - White noise
  - Random walk
  - Stationary Gaussian process

Random Process and stationary (1)

A random process X_t is normally assumed to be composed of a finite number of random variables $\{X_{t_1}, X_{t_2}, \dots, X_{t_n}\}$ with the distribution as,

- $F_{X_{t_1},X_{t_2},\dots,X_{t_n}}(x_1, x_2, \dots, x_n) = P(X_{t_1} \le x_1, X_{t_2} \le x_2, \dots, X_{t_n} \le x_n)$
- where $x_i, i = 1{,}2.,, n$ are real numbers
- The random variables $\{X_{t_1}, X_{t_2}, \dots, X_{t_n}\}$ form a random process
- The values $(x_1, x_2, \dots, x_n)$ is called a realization from the random process X_t, such as one measurement sample from the X_t

Random Process and stationary (2)

- The random process X_t is said to be first order (strictly/strongly) stationary if
  - ✓The marginal distribution $F_{X_{t_i}}(x) = F_{X_{t_j}}(x)$ for any X_{t_i}, X_{t_j} in the process
- The random process is said to be second order (strictly/strongly) stationary if
  - ✓The marginal distribution $F_{X_{t_i}, X_{t_j}}(x_1, x_2) = F_{X_{t_{i+k}}, X_{t_{j+k}}}(x_1, x_2)$
- The high order stationary process can be defined in the same manner

Properties of stationary RP

- The mean and variance function of the random process X_t
  - ✓ $\mu_t = E[X_t]$
  - ✓ $\sigma_t = E[(X_t - \mu_t)^2]$
- The covariance and correlation between X_{t_1} and X_{t_2}
  - ✓ $\gamma(t_1, t_2) = E\left[(X_{t_1} - \mu_{t_1})(X_{t_2} - \mu_{t_2})\right]$
  - ✓ $\rho(t_1, t_2) = \frac{\gamma(t_1, t_2)}{\sigma_{t_1}\sigma_{t_2}}$
- For strictly stationary random process, for any *k*, we have the following
  - ✓ $\gamma(t_1, t_2) = \gamma(t_1 + k, t_2 + k)$ and
  - ✓ $\rho(t_1, t_2) = \rho(t_1 + k, t_2 + k)$
- It means that for a stationary random process X_t with finite first two moments, the covariance and correlation only depends on the time difference, i.e.,
  - ✓ $\gamma(k) = \gamma(t, t + k)$
  - ✓ $\rho(k) = \rho(t, t + k)$

Strict and weak stationary RP against Gaussian random process

- A process to be 2nd order weakly stationary if
  - If $\mu_t = E[X_t]$ is independent of *t*
  - The correlation/covariance structure $\gamma(k) = \gamma(t, t+k)$ is independent of *t*
- *Normally, a strictly stationary process also fulfil the above first 2 order relationships*
- *But a weakly stationary cannot → strictly stationary as* $F_{X_{t_i}, X_{t_j}}(x_1, x_2) = F_{X_{t_{i+k}}, X_{t_{j+k}}}(x_1, x_2)$
- *However, since a Gaussian process is uniquely determined by the first 2 moments, i.e.,*

$$f_{\mathbf{X}}(x_1, \ldots, x_k) = \frac{\exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})\right)}{\sqrt{(2\pi)^k |\boldsymbol{\Sigma}|}}$$

- *Therefore, for a Gaussian process, the first 2 order weakly stationary also means strictly stationary. This is why Gaussian process has been used in the time series analysis.*

Stationary process

Autocorrelation function (ACF) of a stationary X_t is defined as,

- $\gamma(k) = \gamma(t, t+k) = Cov(X_t, X_{t+k}) = E[(X_t - \mu)(X_{t+k} - \mu)]$
- $\gamma(0) = \gamma(t, t) = Cov(X_t, X_t)$
- The mean of X_t is independent of time, *i.e.*, μ_t is a constant
- $\gamma(h) = \gamma(-h)$

Without mention specifically, stationary means a 2nd order weakly stationary process, because it is difficult to assess the strict stationarity of a process from a single dataset (realization).

Outline

- Different random data types
- Properties to describe random process
  - Mean, variance
  - Covariance and correlation function
  - Stationary random process
  - Weakly stationary random process
- **Typical examples of random process**
  - White noise
  - Random walk
  - Stationary Gaussian process

A white noise process

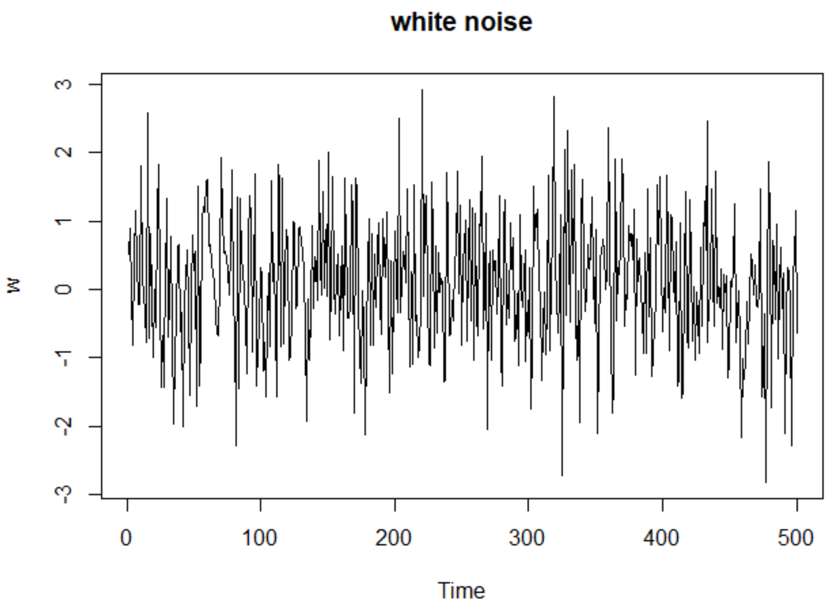
A process W_t is defined as white noise such that the random variable at any time of the process is i.i.d. with zero mean and constant variance

✓$W_t \sim iid(0, \sigma_w^2)$ for any $t = t_1, t_2, \ldots, t_n$ in the process

✓If the random variables are following **Gaussian** distribution, it is ***Gaussian white noise***, that are most common used in the engineering applications

Its stationarity as derived by:

✓$\gamma(h) = cov(W_t, W_{t+h}) = \begin{cases} \sigma_w^2, h = 0 \\ 0, h \neq 0 \end{cases}$

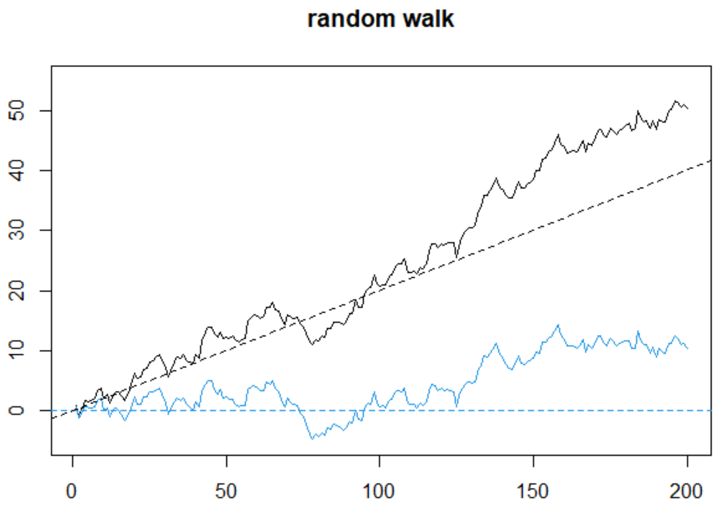


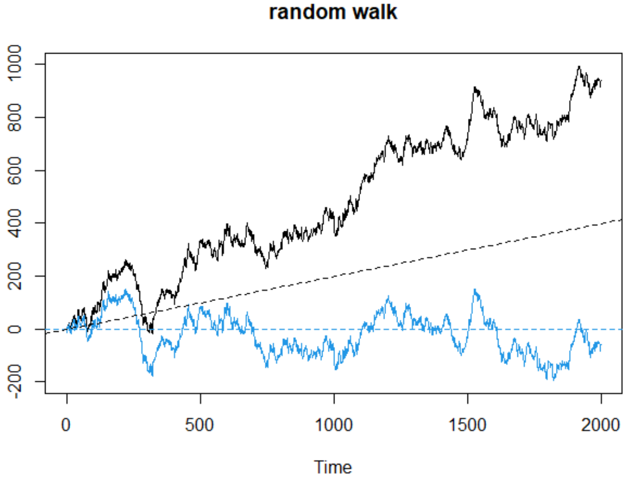
A random walk

A random walk process can be written in two ways

- ✓ $X_t = \delta + X_{t-1} + W_t$ (AR)
- ✓ $X_t = \delta t + \sum_{j=1}^{t} W_j$ (MA)
- ✓ If $\delta = 0$, X_t is a simple random walk, while $\delta \neq 0$, X_t is a random walk with shift

- For a random walk with shift $\delta \neq 0$, its mean values is not constant over time. It is **not** a stationary process
- For the (simple) random walk process X_t, its autocovariance can be derived as
  - ✓ $\gamma(s,t) = cov(X_s, X_t) = cov\left(\sum_{j=1}^{s} W_j, \sum_{k=1}^{t} W_k\right) = min\{s,t\}\sigma_w^2$
  - ✓ The covariance is dependent on the time *t*





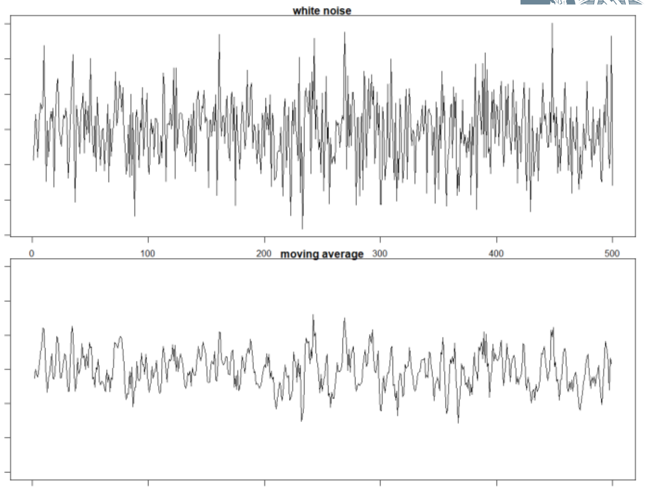
Moving average process

- Let a white noise process denote by $W_t \sim iid(0, \sigma_W^2)$
- A moving average process to smooth the noise can be defined as: $X_t = \frac{1}{3}(W_{t-1} + W_t + W_{t+1})$
- The covariance function defined $\gamma_X(s,t) = cov(X_s, X_t)$
- When *s*=*t*, it can be derived as

$$\begin{aligned}\gamma_v(t,t) &= \tfrac{1}{9}\text{cov}\{(w_{t-1} + w_t + w_{t+1}), (w_{t-1} + w_t + w_{t+1})\} \\ &= \tfrac{1}{9}[\text{cov}(w_{t-1}, w_{t-1}) + \text{cov}(w_t, w_t) + \text{cov}(w_{t+1}, w_{t+1})] \\ &= \tfrac{3}{9}\sigma_w^2.\end{aligned}$$

- When *s*=*t+1*, it can be derived as

$$\begin{aligned}\gamma_v(t+1,t) &= \tfrac{1}{9}\text{cov}\{(w_t + w_{t+1} + w_{t+2}), (w_{t-1} + w_t + w_{t+1})\} \\ &= \tfrac{1}{9}[\text{cov}(w_t, w_t) + \text{cov}(w_{t+1}, w_{t+1})] \\ &= \tfrac{2}{9}\sigma_w^2,\end{aligned}$$



$$\gamma_X(s,t) = \begin{cases} \frac{3}{9}\sigma_w^2 & s = t, \\ \frac{2}{9}\sigma_w^2 & |s-t| = 1, \\ \frac{1}{9}\sigma_w^2 & |s-t| = 2, \\ 0 & |s-t| > 2. \end{cases}$$

Q: is a MV process stationary?

Trend stationarity

A trend stationary process is defined as follows.

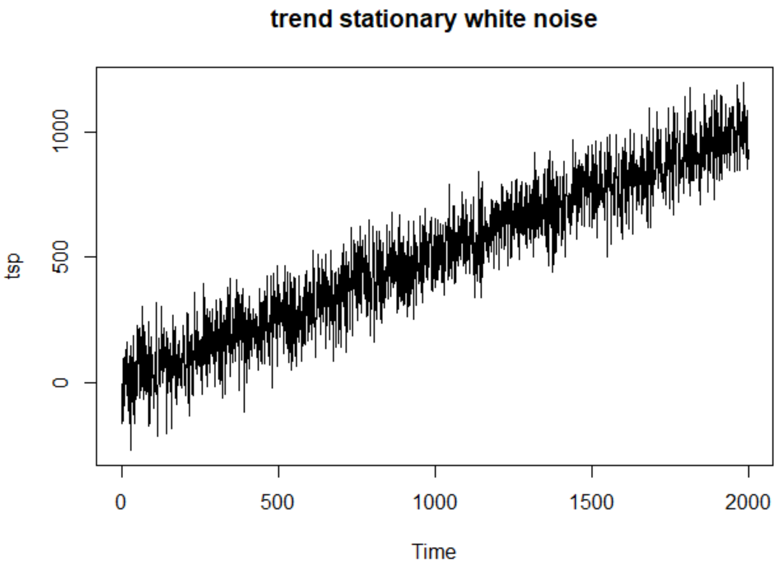
For $X_t = \alpha + \beta t + Y_t$ (Y_t: a stationary process)

Mean of X_t is $\mu_{X_t} = E[X_t] = \alpha + \beta t + \mu_Y$

The autocovariance function can be derived as,

$$\gamma_X(h) = cov(X_{t+h}, X_t) = E\left[\left(X_{t+h} - \mu_{X_{t+h}}\right)\left(X_t - \mu_{X_t}\right)\right] = E[(Y_{t+h} - \mu_Y)(Y_t - \mu_Y)] = \gamma_Y(h)$$

- ✓Mean is dependent of time
- ✓The autocovariance is independent of time
- ✓It is called the trend stationary, i.e., if we take away the trend, the residual can be regarded as a stationary process

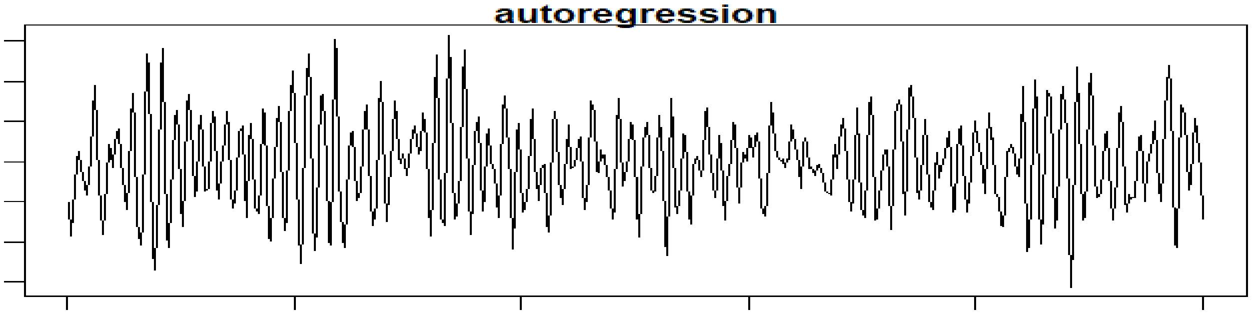


This property will be important for the model exploration of time series data.

Autoregressive (AR) process/model of various order (normally 1^{st} of 2^{nd})

Autoregressive (AR) process/model of various order (normally 1st of 2nd)

- Let write an AR process in terms of white noise process as:
  - ✓ $X_t = X_{t-1} - 0.9X_{t-2} + W_t$
  - ✓ Itself represents a regression/prediction of current value in terms of two previous values + noise
- ***The big issue in AR models is the initial values.***
- In the application function “filter” of a stationary white noise in R, the initial values are often set to be $X_1 = W_1$, $X_2 = X_1 + W_1$, which do not satisfy the AR model as above
- But an easy fix for practical applications is to get longer time series and remove initial values at the beginning of the time series.



If we assume that we know the initial values of a AR process, is the AR process a stationary process?

Two (or more) time series

Two time series denote by X_t, Y_t, are said to be jointly stationary if

- ✓ X_t and Y_t *are each stationary*
- ✓The cross-covariance function is a function of only the time lag h *as:*

$$\gamma_{X,Y}(h) = cov(X_{t+h}, Y_t) = E[(X_{t+h} - \mu_X)(Y_t - \mu_Y)]$$

The cross-correlation function of the jointly stationary time series X_t, Y_t, is defined as,

- ✓$\rho_{X,Y} = \frac{\gamma_{XY}(h)}{\sqrt{\gamma_X(0)\gamma_Y(0)}}$
- ✓$-1 \leq \rho_{X,Y} \leq 1$
- ✓$\rho_{X,Y}$ *is normally NOT symmetric, i.e.,* $\rho_{X,Y}(h) \neq \rho_{X,Y}(-h)$

One example of the jointly stationary time series through a white noise process $W_t \sim \mathrm{iid}(0, \sigma_W^2)$

- ✓$X_t = W_t + W_{t-1}$ *and* $Y_t = W_t - W_{t-1}$
- ✓$\gamma_X(0) = \gamma_Y(0) = 2\sigma_W^2$, $\gamma_X(1) = \gamma_X(-1) = \sigma_W^2$, $\gamma_Y(1) = \gamma_Y(-1) = -\sigma_W^2$
- ✓$\gamma_{X,Y}(1) = cov(X_{t+1}, Y_t) = cov(W_{t+1} + W_t, W_t + W_{t-1}) = \sigma_W^2$, $\gamma_{X,Y}(0) = 0$, $\gamma_{X,Y}(-1) = -\sigma_W^2$
- ✓The autocovariance and cross-covariance functions depend only on lag h → joint stationary

$$\rho_{xy}(h) = \begin{cases} 0 & h = 0, \\ 1/2 & h = 1, \\ -1/2 & h = -1, \\ 0 & |h| \geq 2. \end{cases}$$

Estimation of (auto) correlation (ACF)

For a random process X_t, there is often one realization/measurement (time series), i.e., we have only one sample data for one RV $X_i, i = 1,2, \ldots, t$.

- ✓However, for estimation of correlation between two RVs, we will need several data points
- ✓If the process X_t is a stationary random process, we can assume/treat that all the data points can be associated with each of the RV $X_i, i = 1,2, \ldots, t$

For the stationary random process X_t, we have

- ✓The mean of all RVs in the process is constant, i.e., $\mu_t = \mu$
- ✓The sample (time series) mean is defined as: $\bar{X} = \frac{1}{n} \sum_{t=1}^{n} X_t$
- ✓The expected value of the sample mean: $E[\bar{X}] = \mu$
- ✓The standard error (variance) of the sample mean is estimated by

$$\begin{aligned}
\operatorname{var}(\bar{x}) &= \operatorname{var}\left(\frac{1}{n}\sum_{t=1}^{n} x_t\right) = \frac{1}{n^2}\operatorname{cov}\left(\sum_{t=1}^{n} x_t, \sum_{s=1}^{n} x_s\right) \\
&= \frac{1}{n^2}\Big(n\gamma_x(0) + (n-1)\gamma_x(1) + (n-2)\gamma_x(2) + \cdots + \gamma_x(n-1) \\
&\qquad + (n-1)\gamma_x(-1) + (n-2)\gamma_x(-2) + \cdots + \gamma_x(1-n)\Big) \\
&= \frac{1}{n}\sum_{h=-n}^{n}\left(1 - \frac{|h|}{n}\right)\gamma_x(h).
\end{aligned}$$

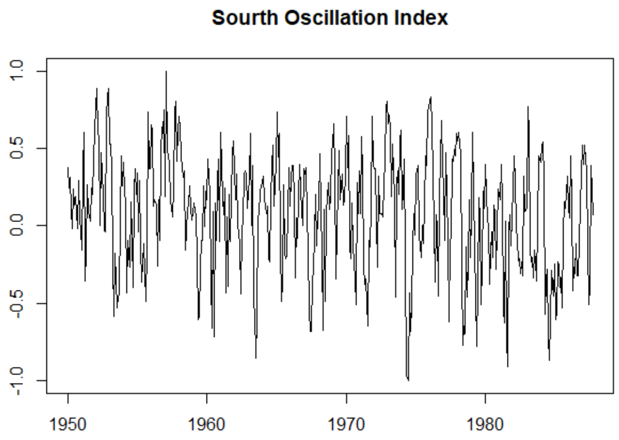
Estimation of sample ACF

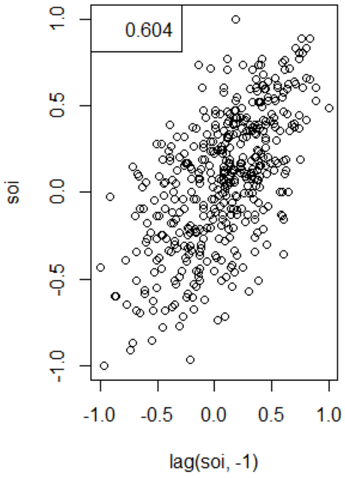
The sample autocovariance function of a stationary process X_t can be defined as

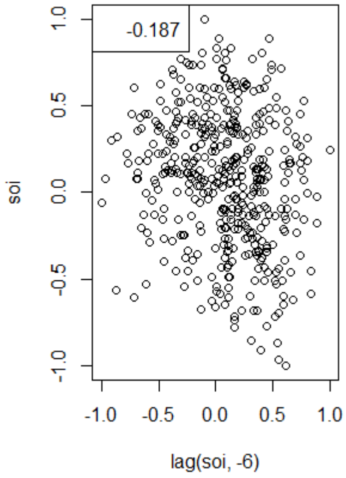
- ✓ $\hat{\gamma}(h) = E[(X_{t+h} - \bar{X})(X_t - \bar{X})]$

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

- ✓ $\hat{\gamma}(h) = \hat{\gamma}(-h)$ for h=0,1,…,n-1
- ✓ The sample autocorrelation function is defined as $\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}$



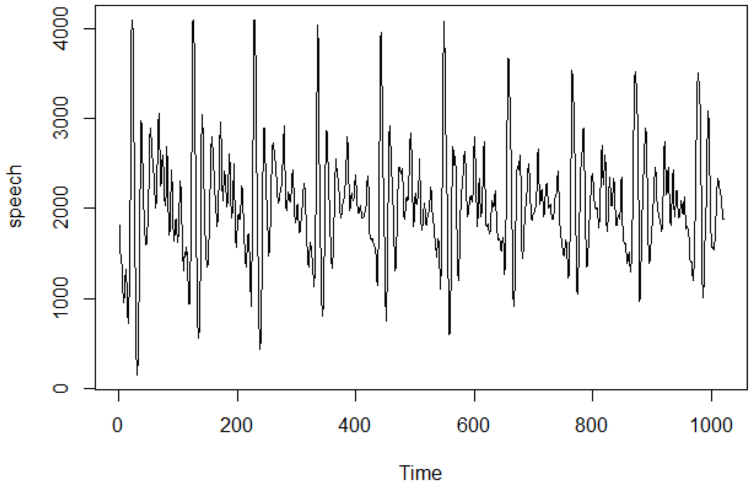


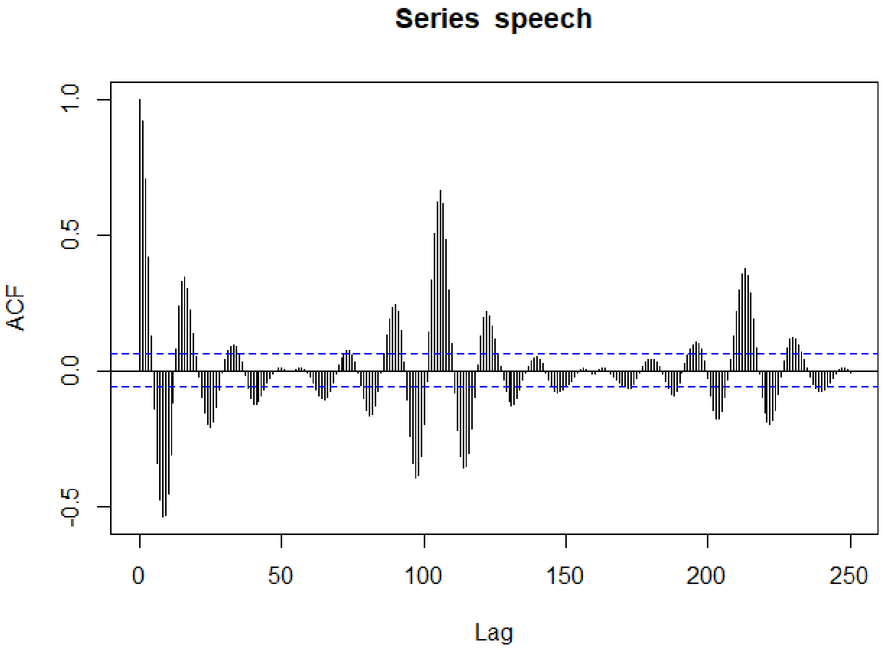


Similarly, for a two process X_t, Y_t, the sample cross-correlation function is computed by:

- ✓ $\hat{\gamma}_{xy}(h) = E[(X_{t+h} - \bar{X})(Y_t - \bar{Y})] = \frac{1}{n}\sum_{t=1}^{n-h}(x_{t+h} - \bar{x})(y_t - \bar{y})$
- ✓ $\hat{\rho}_{xy}(h) = \frac{\hat{\gamma}_{xy}(h)}{\sqrt{\hat{\gamma}_x(0)\hat{\gamma}_y(0)}}$

An example: ACF of a speech signal

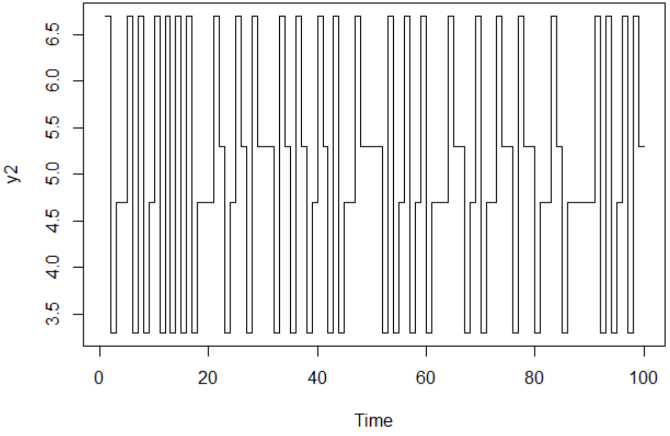




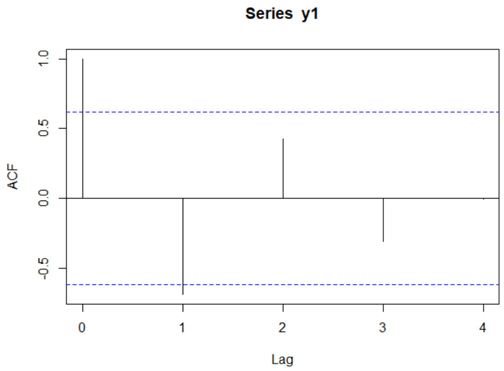
Run command from R: acf(speech,250)

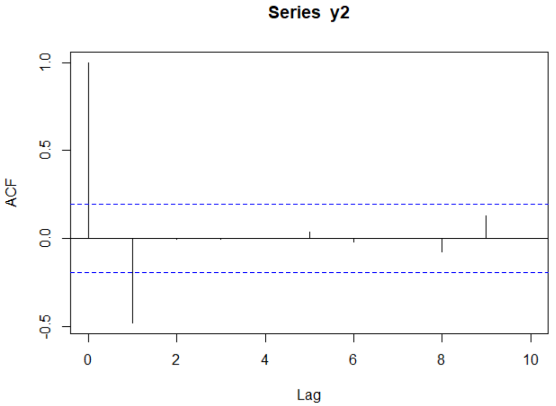
ACF: a simulated sample

- Simulate the time series process with $y_t = 5 + x_t - 0.7x_{t-1}$
  - X_t represent the head or tail when tossing a fair
  - X_t=1 → head; X_t=-1 → tail
- Run the following code in R to simulate the time series
  - set.seed(101010)
  - x1 = 2*rbinom(11, 1, .5) - 1 # simulated sequence of coin tosses
  - x2 = 2*rbinom(101, 1, .5) - 1
  - y1 = 5 + filter(x1, sides=1, filter=c(1,-.7))[-1]
  - y2 = 5 + filter(x2, sides=1, filter=c(1,-.7))[-1]
  - plot.ts(y1, type='s'); plot.ts(y2, type='s') # plot both series (not shown)
- The results we can get




```
> c(mean(y1), mean(y2))
[1] 5.080 5.002
> acf(y1, lag.max=4, plot=FALSE)

Autocorrelations of series ‘y1’, by lag

     0      1      2      3      4
 1.000 -0.688  0.425 -0.306 -0.007
> acf(y2, lag.max=4, plot=FALSE)

Autocorrelations of series ‘y2’, by lag

     0      1      2      3      4
 1.000 -0.480 -0.002 -0.004  0.000
```





AVANCEZ
1829
CHALMERS
UNIVERSITY OF TECHNOLOGY