# Stationarity & introductory functions

## FISH 507 – Applied Time Series Analysis

Mark Scheuerell

10 Jan 2019

# Topics for today

Characteristics of time series

- Expectation, mean & variance
- Covariance & correlation
- Stationarity
- Autocovariance & autocorrelation
- Correlograms

White noise

Random walks

Backshift & difference operators

# Expectation & the mean

The expectation ($E$) of a variable is its mean value in the population

$\mathrm{E}(x) \equiv$ mean of $x = \mu$

We can estimate $\mu$ from a sample as

$$m = \frac{\sum_{i=1}^{N} x_i}{N}$$

# Variance

$\mathrm{E}([x-\mu]^2) \equiv$ expected deviations of $x$ about $\mu$

$\mathrm{E}([x-\mu]^2) \equiv$ variance of $x = \sigma^2$

We can estimate $\sigma^2$ from a sample as

$$s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - m)^2$$

# Covariance

If we have two variables, $x$ and $y$, we can generalize variance

$$\sigma^2 = \mathrm{E}([x_i - \mu][x_i - \mu])$$

into *covariance*

$$\gamma_{x,y} = \mathrm{E}([x_i - \mu_x][y_i - \mu_y])$$

# Covariance

If we have two variables, $x$ and $y$, we can generalize variance

$$\sigma^2 = \mathrm{E}([x_i - \mu][x_i - \mu])$$

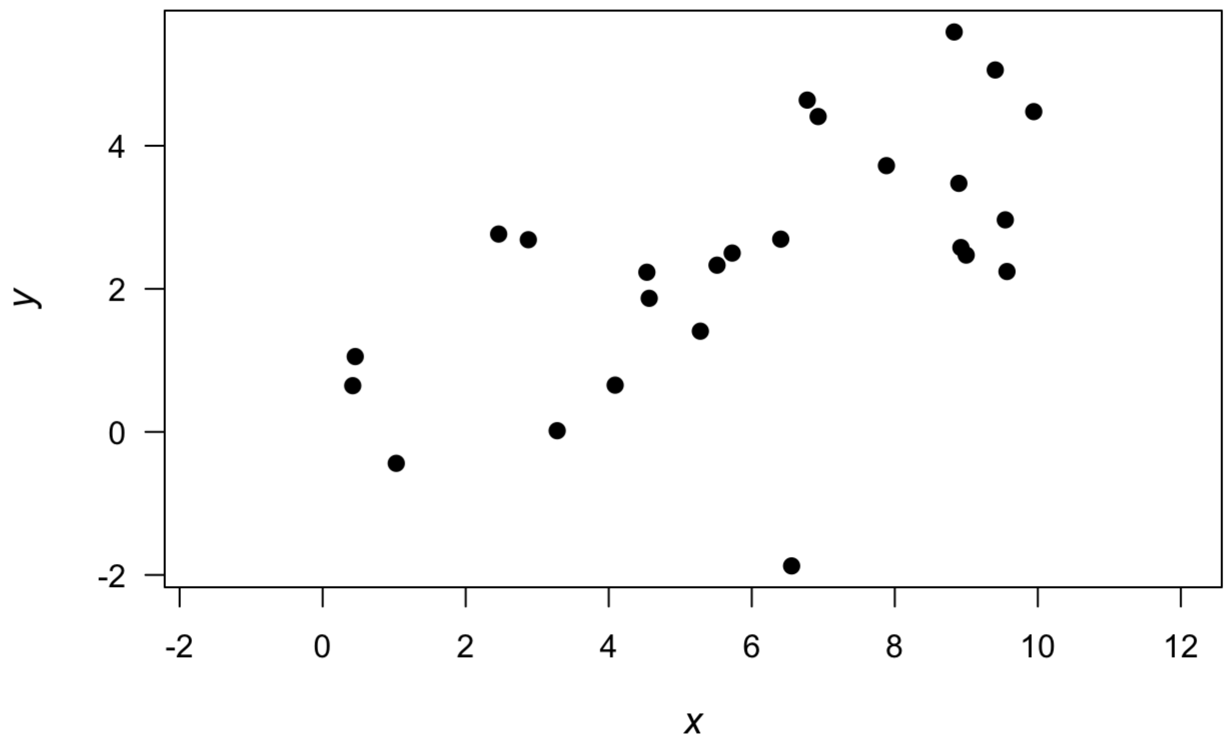
into *covariance*

$$\gamma_{x,y} = \mathrm{E}([x_i - \mu_x][y_i - \mu_y])$$

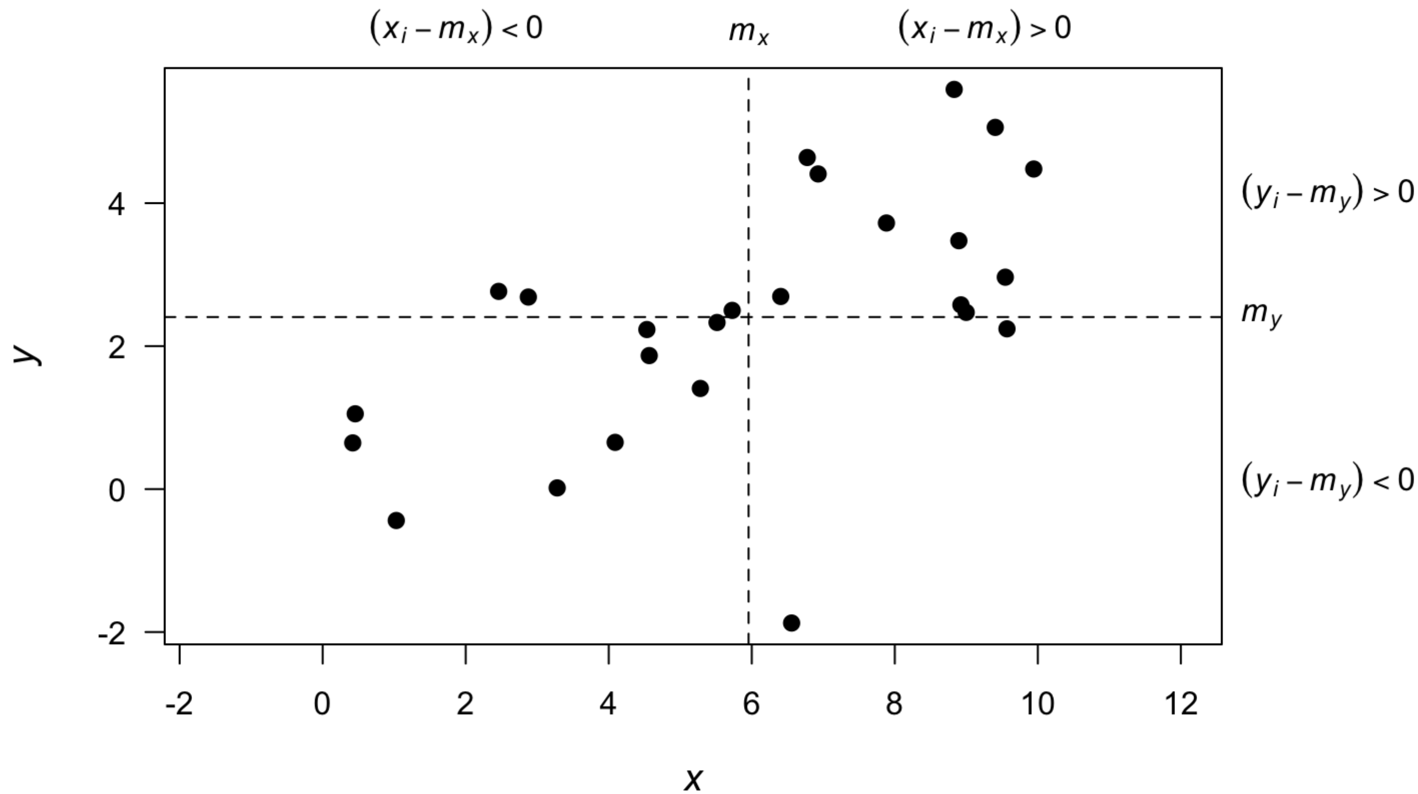
We can estimate $\gamma_{x,y}$ from a sample as

$$\mathrm{Cov}(x, y) = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - m_x)(y_i - m_y)$$

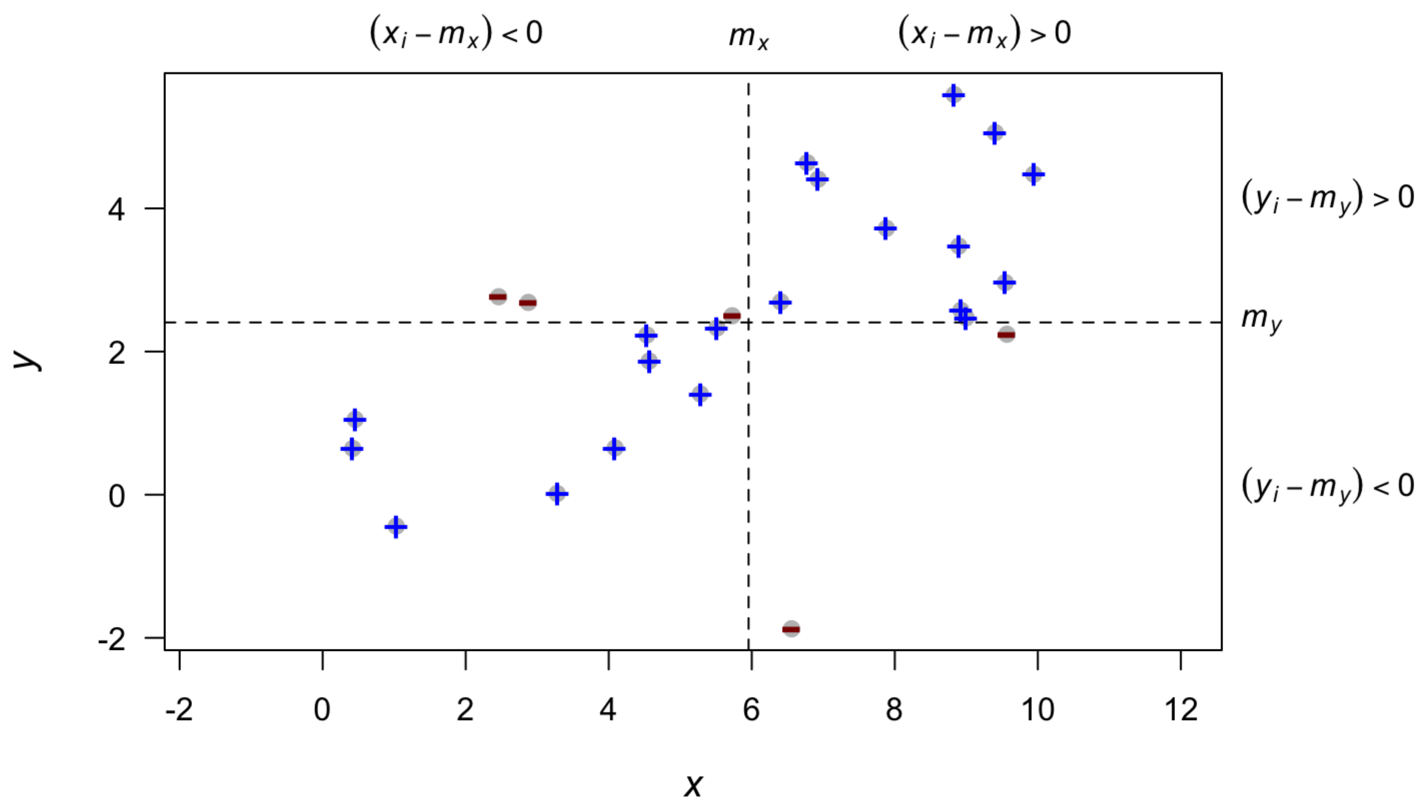
# Graphical example of covariance



# Graphical example of covariance



# Graphical example of covariance



# Correlation

*Correlation* is a dimensionless measure of the linear association between 2 variables, $x$ & $y$

It is simply the covariance standardized by the standard deviations

$$\rho_{x,y} = \frac{\gamma_{x,y}}{\sigma_x \sigma_y}$$

$$-1 < \rho_{x,y} < 1$$

# Correlation

*Correlation* is a dimensionless measure of the linear association between 2 variables $x$ & $y$

It is simply the covariance standardized by the standard deviations

$$\rho_{x,y} = \frac{\gamma_{x,y}}{\sigma_x \sigma_y}$$

We can estimate $\rho_{x,y}$ from a sample as

$$\text{Cor}(x, y) = \frac{\text{Cov}(x, y)}{s_x s_y}$$

# Stationarity & the mean

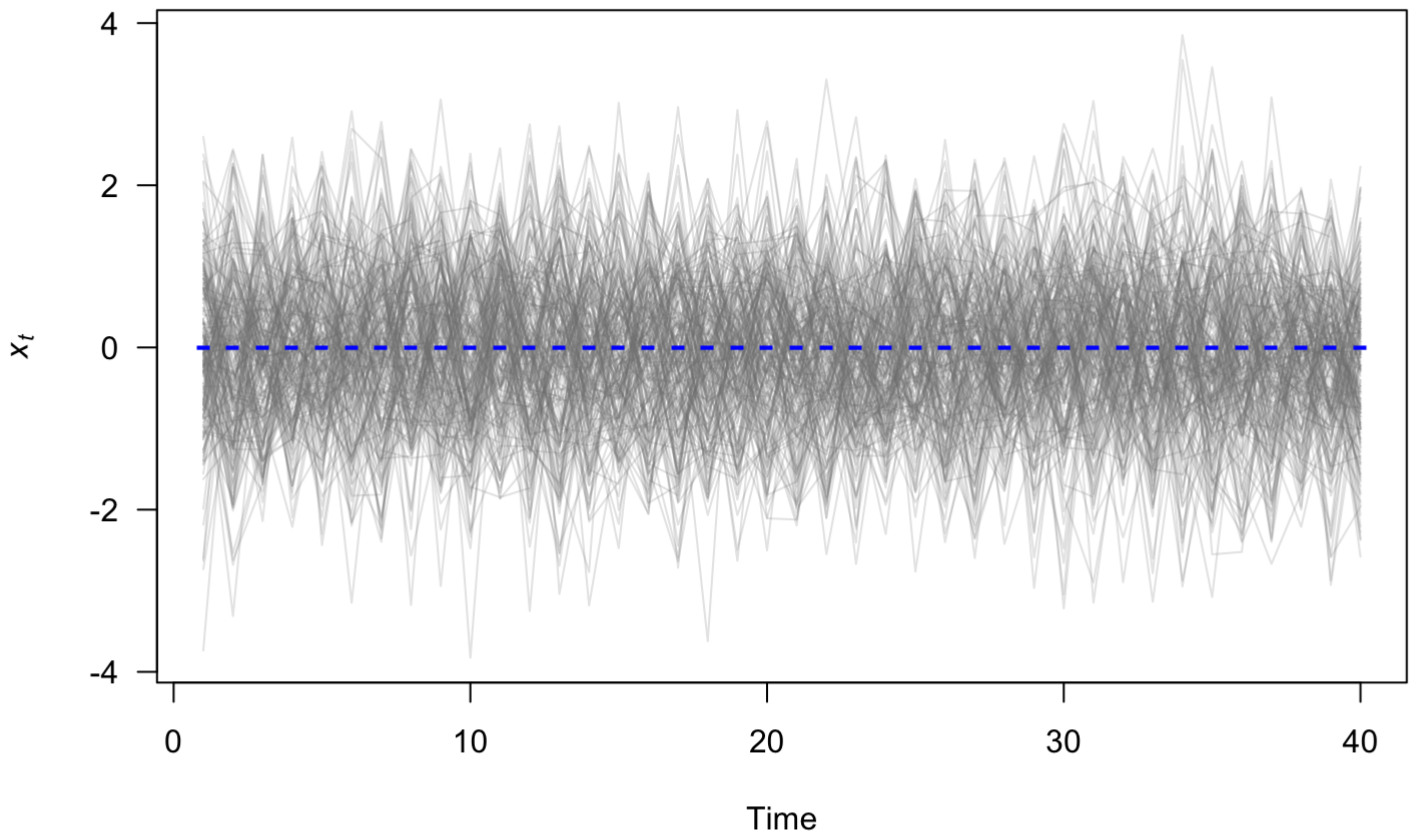
Consider a single value, $x_t$

# Stationarity & the mean

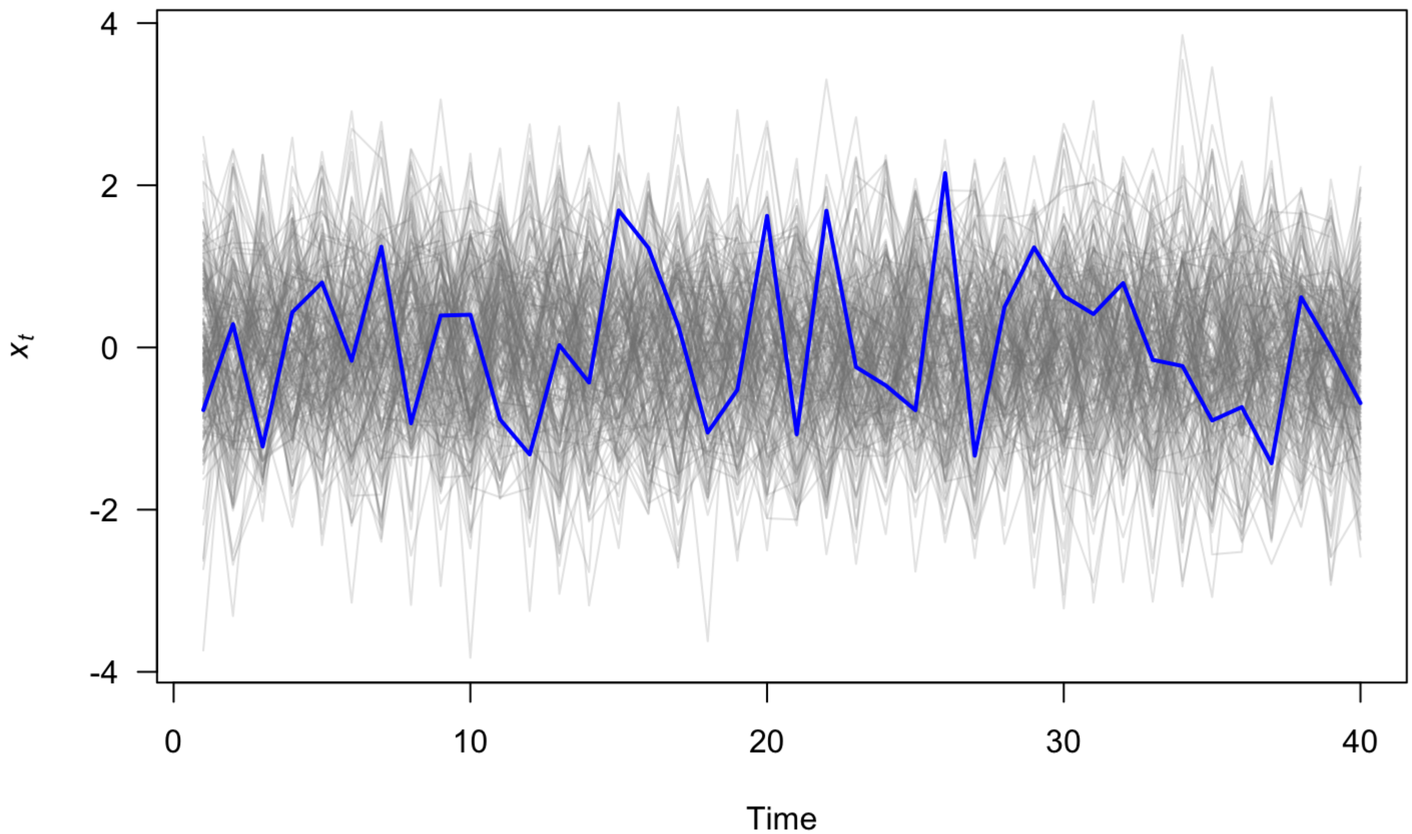
Consider a single value, $x_t$

$\mathrm{E}(x_t)$ is taken across an ensemble of *all* possible time series

# Stationarity & the mean



# Stationarity & the mean




Our single realization is our estimate!

# Stationarity & the mean

If $\mathrm{E}(x_t)$ is constant across time, we say the time series is *stationary* in the mean

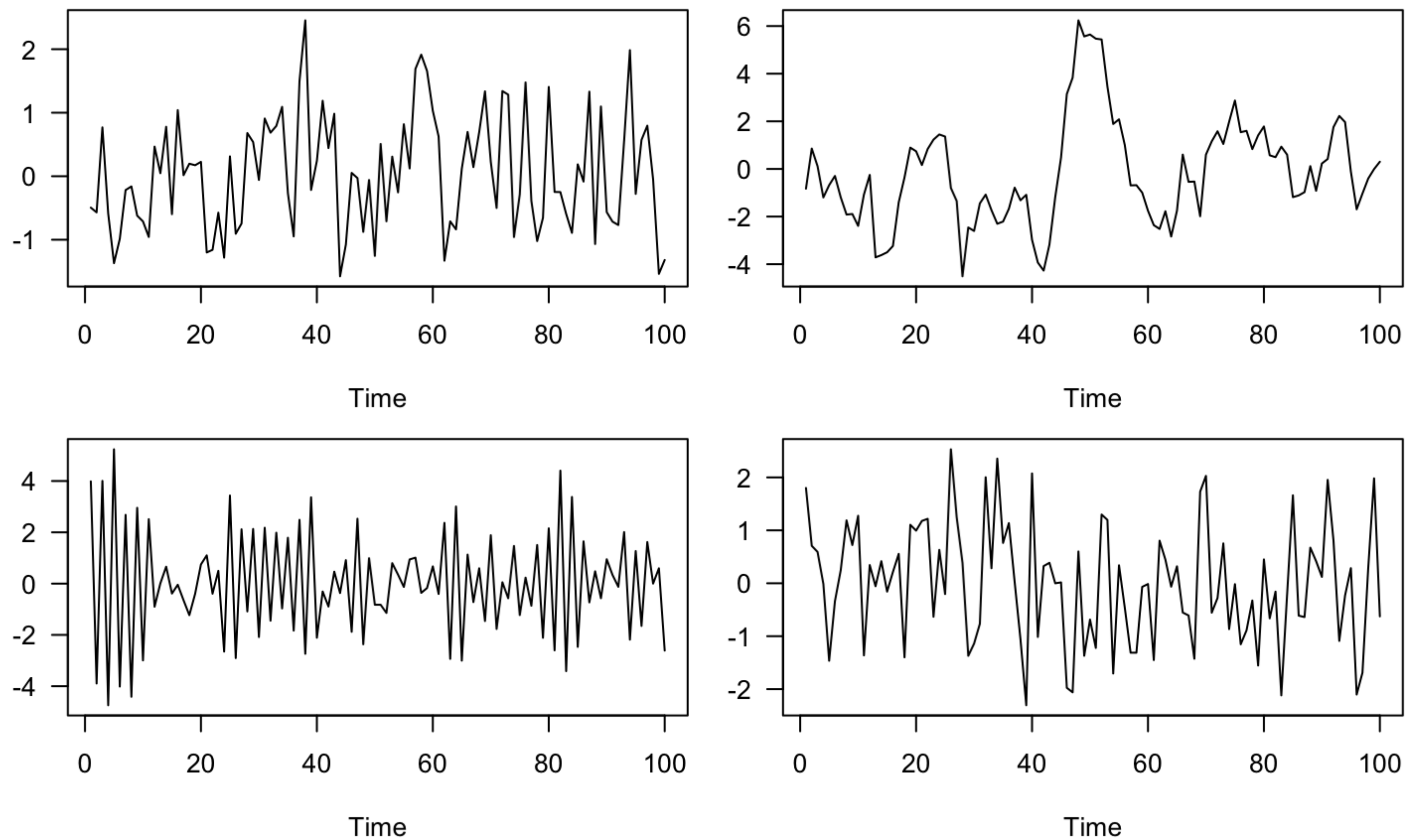
# Stationarity of time series

*Stationarity* is a convenient assumption that allows us to describe the statistical properties of a time series.

In general, a time series is said to be stationary if there is

1. no systematic change in the mean or variance
2. no systematic trend
3. no periodic variations or seasonality

# Identifying stationarity



# Identifying stationarity

Our eyes are really bad at identifying stationarity, so we will learn some tools to help us

# Autocovariance function (ACVF)

For stationary ts, we define the *autocovariance function* ($\gamma_k$) as

$$\gamma_k = \mathrm{E}([x_t - \mu][x_{t+k} - \mu])$$

which means that

$$\gamma_0 = \mathrm{E}([x_t - \mu][x_t - \mu]) = \sigma^2$$

# Autocovariance function (ACVF)

For stationary ts, we define the *autocovariance function* ($\gamma_k$) as

$$\gamma_k = \text{E}([x_t - \mu][x_{t+k} - \mu])$$

"Smooth" series have large ACVF for large $k$

"Choppy" series have ACVF near 0 for small $k$

# Autocovariance function (ACVF)

For stationary ts, we define the *autocovariance function* ($\gamma_k$) as

$$\gamma_k = \mathrm{E}([x_t - \mu][x_{t+k} - \mu])$$

We can estimate $\gamma_k$ from a sample as

$$c_k = \frac{1}{N} \sum_{t=1}^{N-k} (x_t - m)(x_{t+k} - m)$$

# Autocorrelation function (ACF)

The *autocorrelation function* (ACF) is simply the ACVF normalized by the variance

$$\rho_k = \frac{\gamma_k}{\sigma^2} = \frac{\gamma_k}{\gamma_0}$$

The ACF measures the correlation of a time series against a time-shifted version of itself

# Autocorrelation function (ACF)

The *autocorrelation function* (ACF) is simply the ACVF normalized by the variance

$$\rho_k = \frac{\gamma_k}{\sigma^2} = \frac{\gamma_k}{\gamma_0}$$

The ACF measures the correlation of a time series against a time-shifted version of itself

We can estimate ACF from a sample as

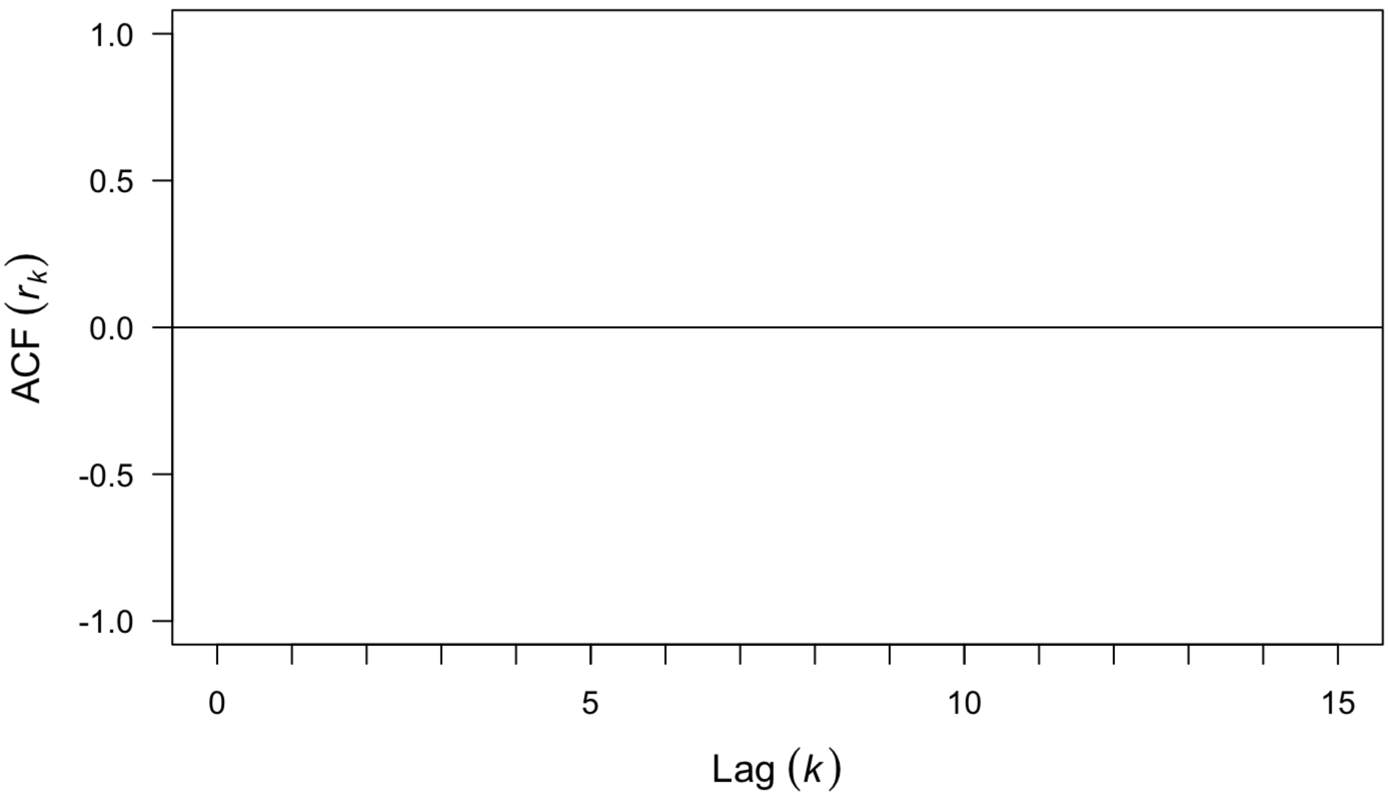
$$r_k = \frac{c_k}{c_0}$$

# Properties of the ACF

The ACF has several important properties:

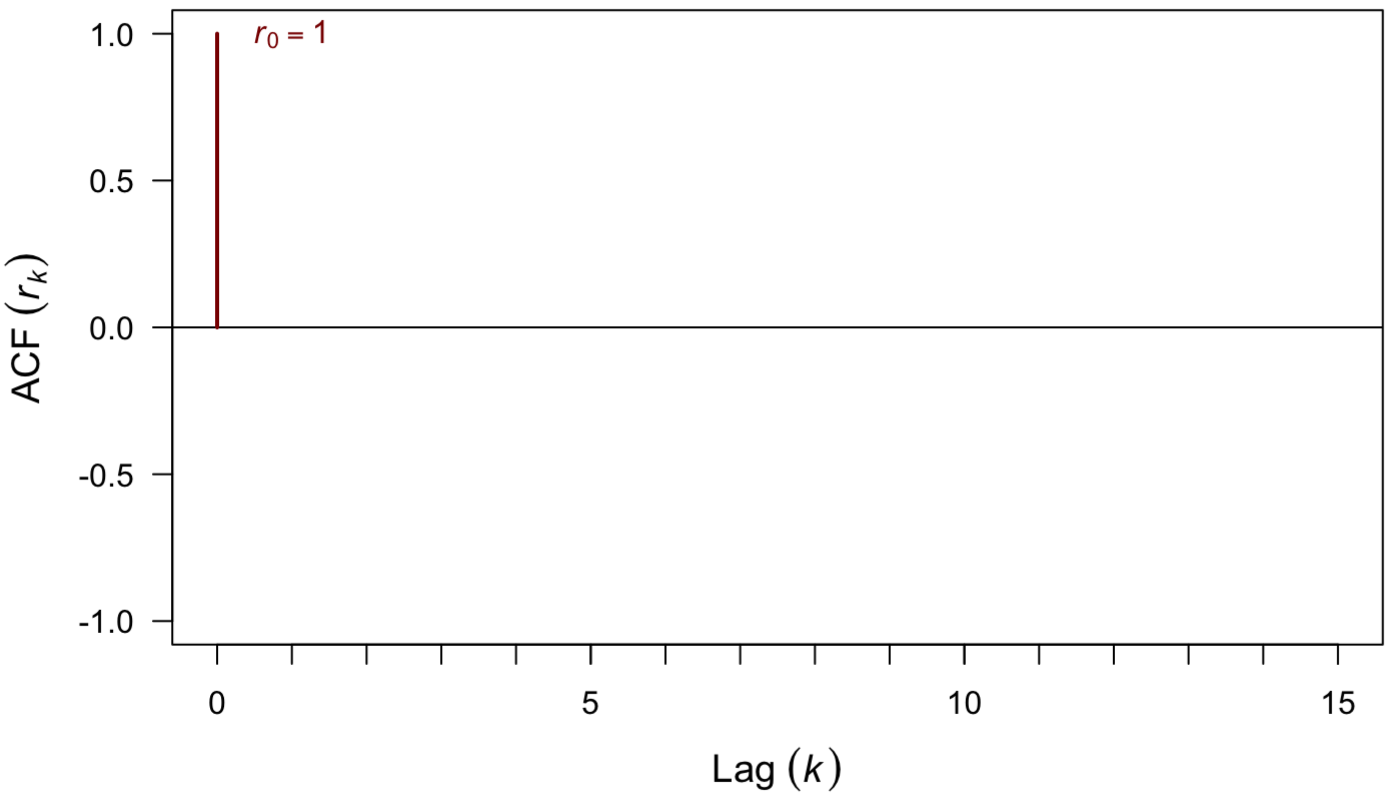
- $-1 \leq r_k \leq 1$
- $r_k = r_{-k}$
- $r_k$ of periodic function is itself periodic
- $r_k$ for the sum of 2 independent variables is the sum of $r_k$ for each of them

# The correlogram



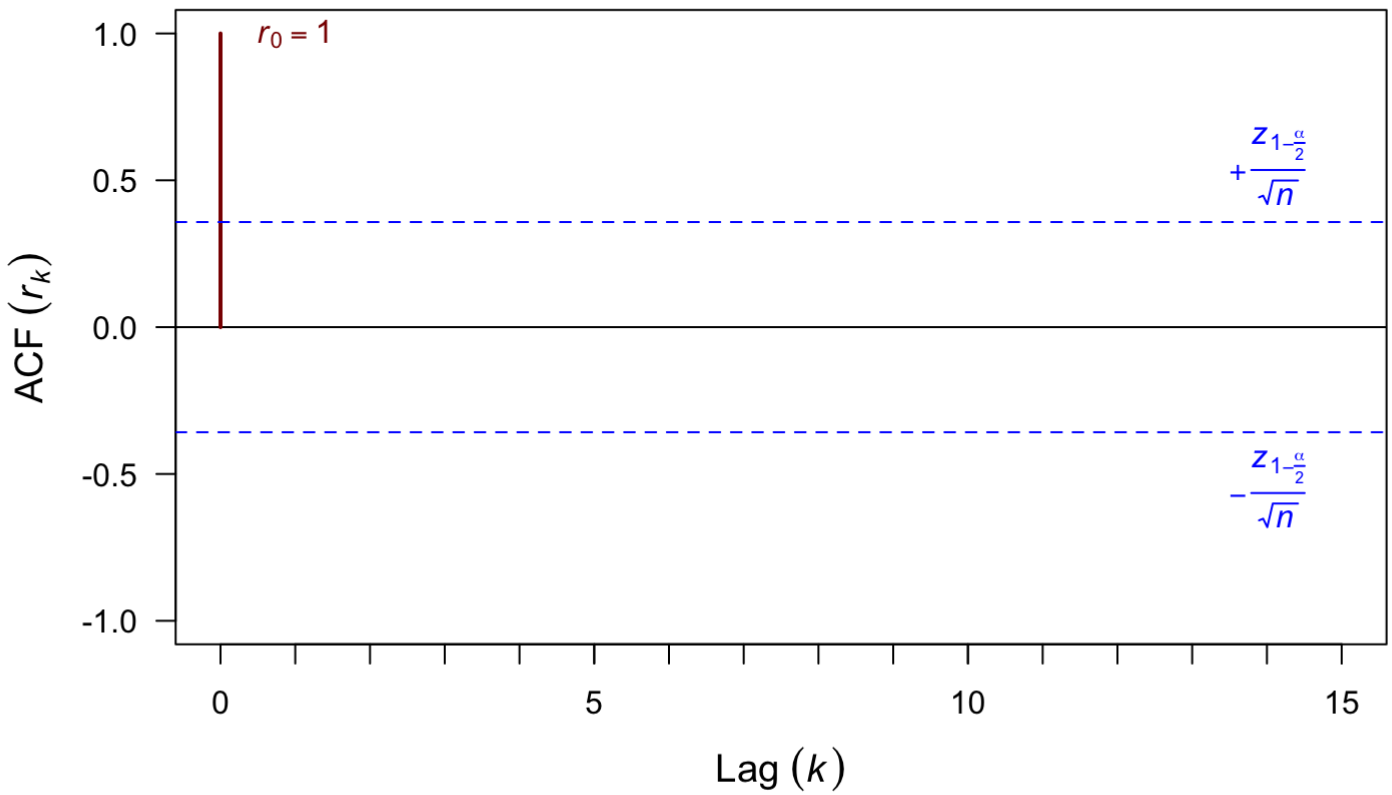

Graphical output for the ACF

# The correlogram



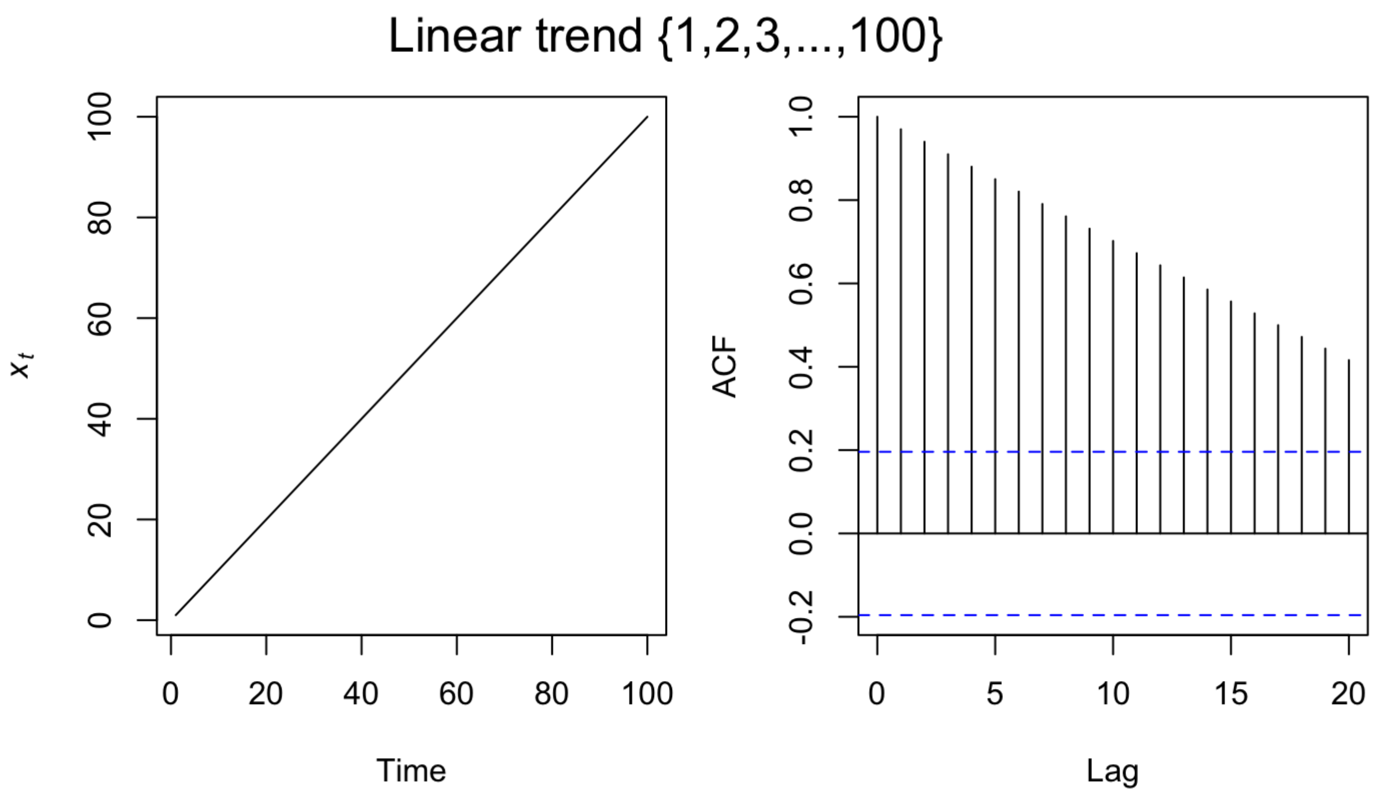

The ACF at lag = 0 is always 1

# The correlogram

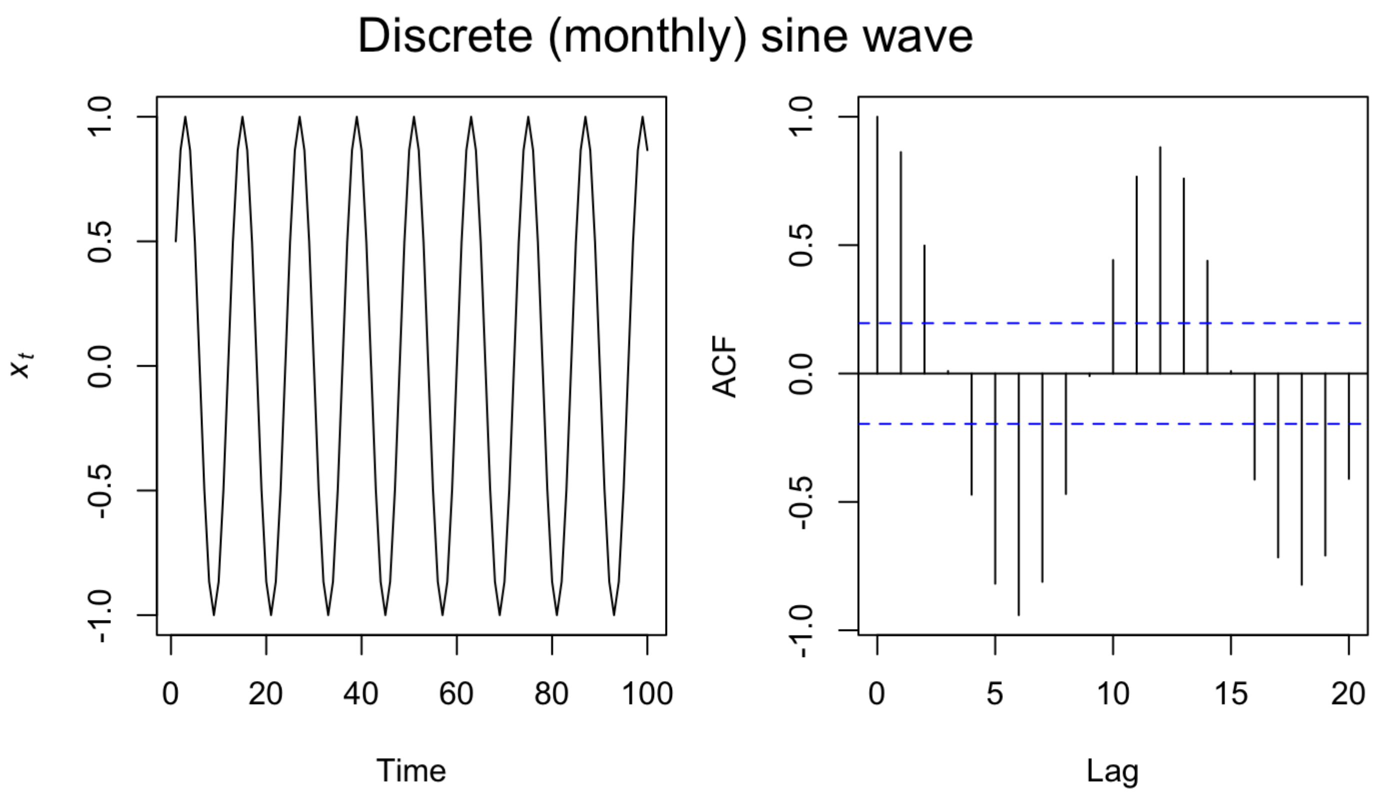



Approximate confidence intervals

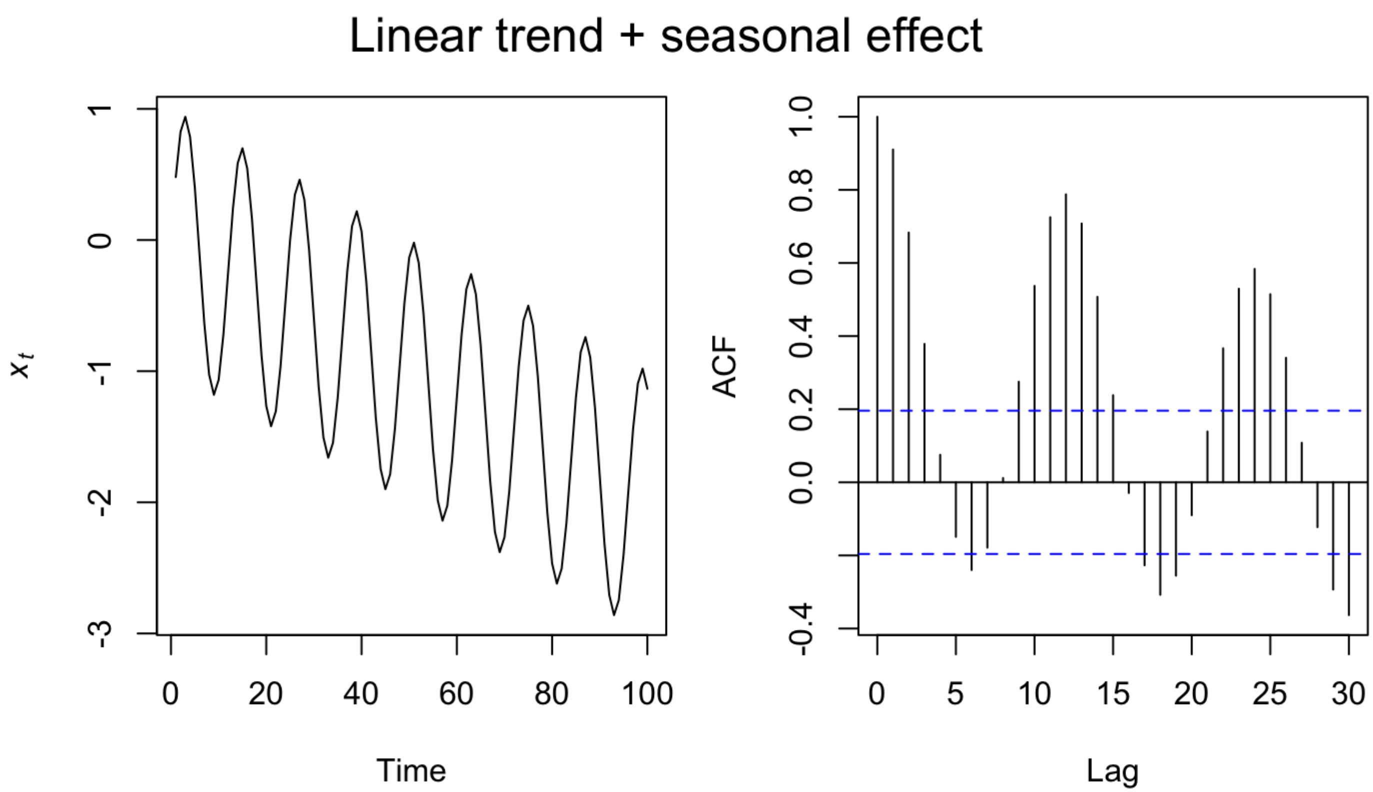
# ACF for deterministic forms



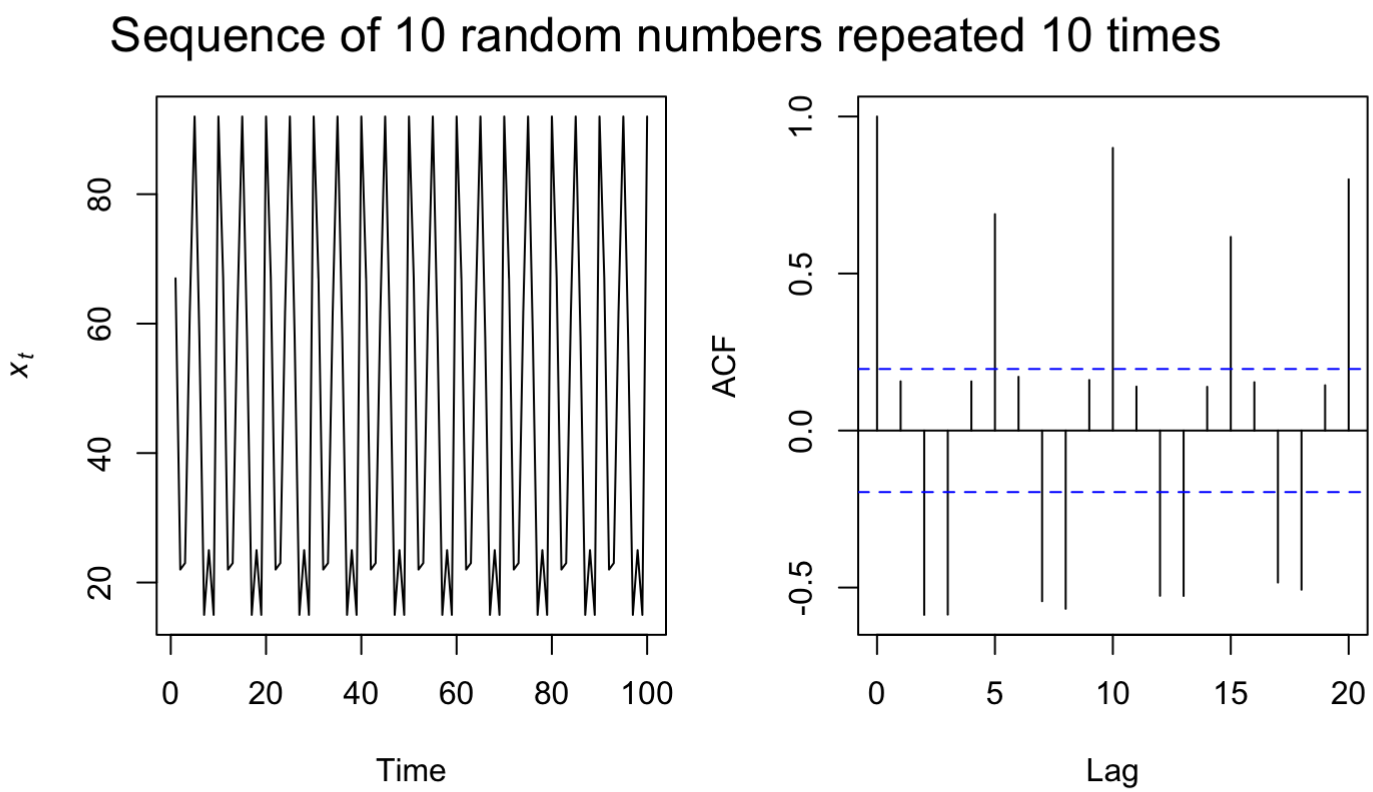
# ACF for deterministic forms



# ACF for deterministic forms



# ACF for deterministic forms



# Induced autocorrelation

Recall the transitive property, whereby

If $A = B$ and $B = C$, then $A = C$

# Induced autocorrelation

Recall the transitive property, whereby

If $A = B$ and $B = C$, then $A = C$

which suggests that

If $x \propto y$ and $y \propto z$, then $x \propto z$

# Induced autocorrelation

Recall the transitive property, whereby

If $A = B$ and $B = C$, then $A = C$

which suggests that

If $x \propto y$ and $y \propto z$, then $x \propto z$

and thus

If $x_t \propto x_{t+1}$ and $x_{t+1} \propto x_{t+2}$, then $x_t \propto x_{t+2}$

# Partial autocorrelation funcion (PACF)

The *partial autocorrelation function* ($\phi_k$) measures the correlation between a series $x_t$ and $x_{t+k}$ with the linear dependence of $\{x_{t-1}, x_{t-2}, \dots, x_{t-k-1}\}$ removed

# Partial autocorrelation funcion (PACF)

The *partial autocorrelation function* ($\phi_k$) measures the correlation between a series $x_t$ and $x_{t+k}$ with the linear dependence of $\{x_{t-1}, x_{t-2}, \dots, x_{t-k-1}\}$ removed

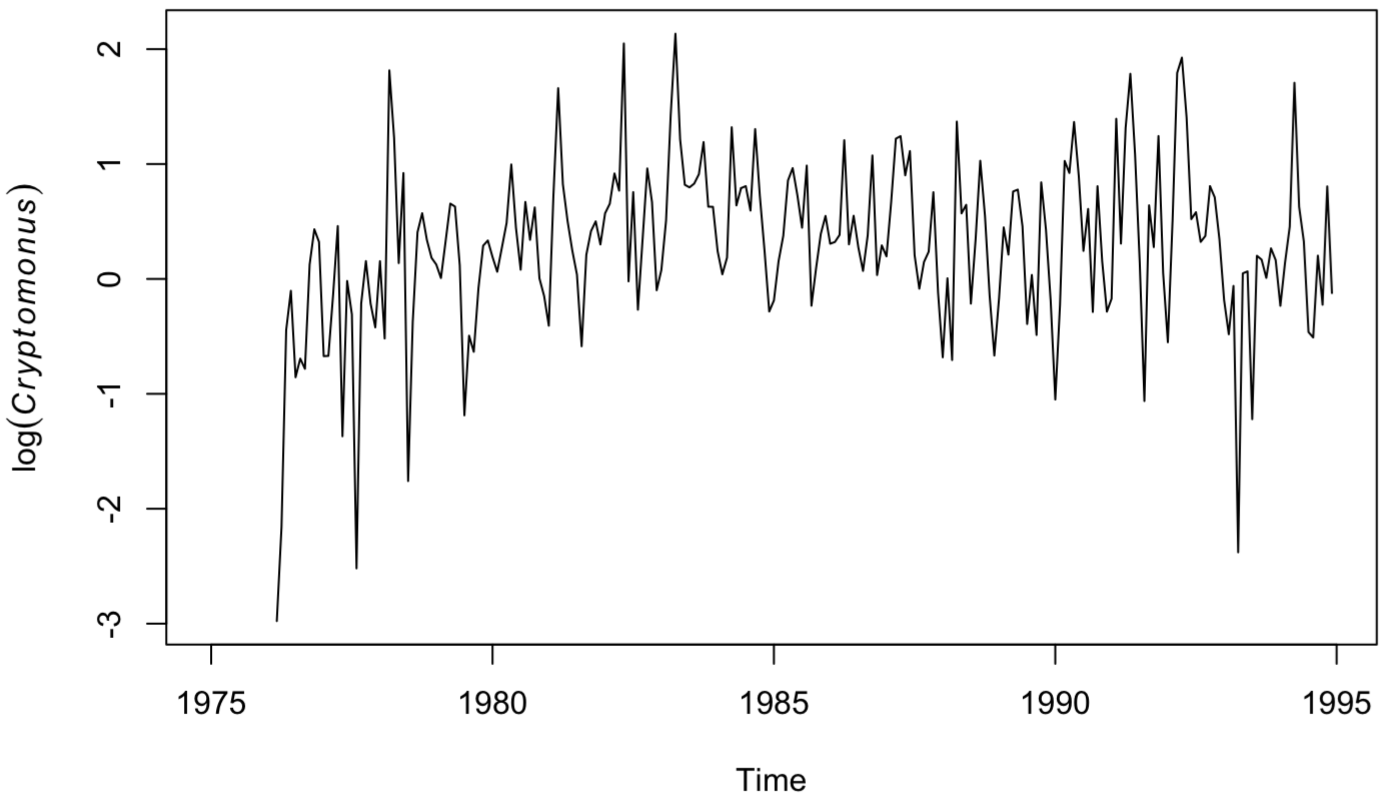
We can estimate $\phi_k$ from a sample as

$$\phi_k = \begin{cases} \text{Cor}(x_1, x_0) = \rho_1 & \text{if } k = 1 \\ \text{Cor}(x_k - x_k^{k-1}, x_0 - x_0^{k-1}) & \text{if } k \geq 2 \end{cases}$$

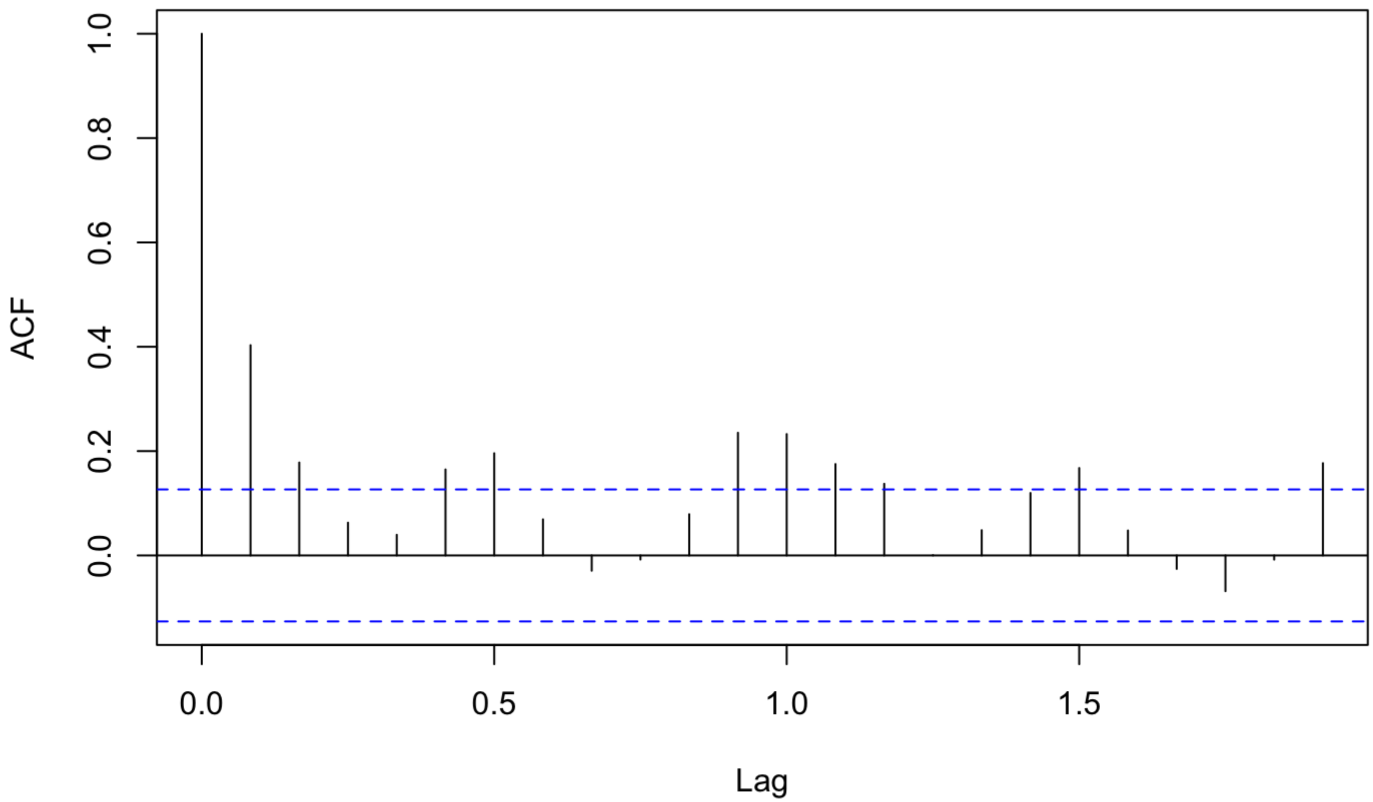
$$x_k^{k-1} = \beta_1 x_{k-1} + \beta_2 x_{k-2} + \cdots + \beta_{k-1} x_1$$

$$x_0^{k-1} = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_{k-1} x_{k-1}$$

# Lake Washington phytoplankton

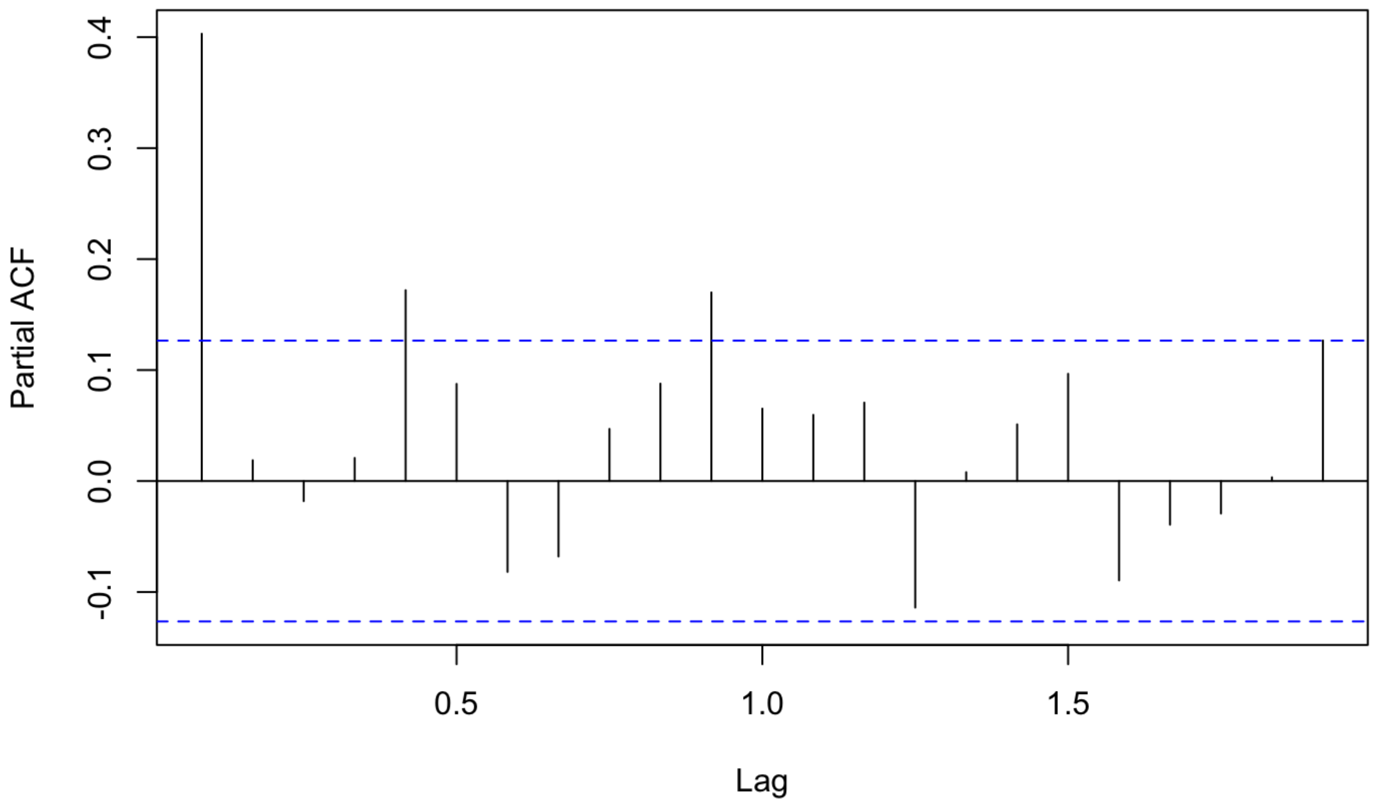


# Lake Washington phytoplankton




Autocorrelation

# Lake Washington phytoplankton




Partial autocorrelation

# ACF & PACF in model selection

The ACF & PACF will be *very* useful for identifying the orders of ARMA models

# Cross-covariance function (CCVF)

Often we want to look for relationships between 2 different time series

We can extend the notion of covariance to *cross-covariance*

# Cross-covariance function (CCVF)

Often we want to look for relationships between 2 different time series

We can extend the notion of covariance to *cross-covariance*

We can estimate $g_k^{x,y}$ from a sample as

$$g_k^{x,y} = \frac{1}{N} \sum_{t=1}^{N-k} (x_t - m_x)(y_{t+k} - m_y)$$

# Cross-correlation function (CCF)

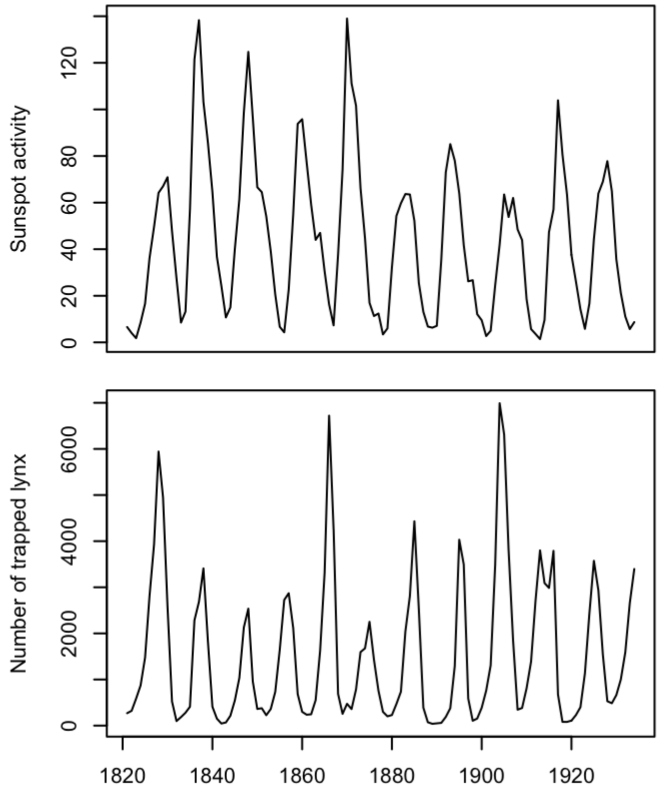
The cross-correlation function is the CCVF normalized by the standard deviations of x & y

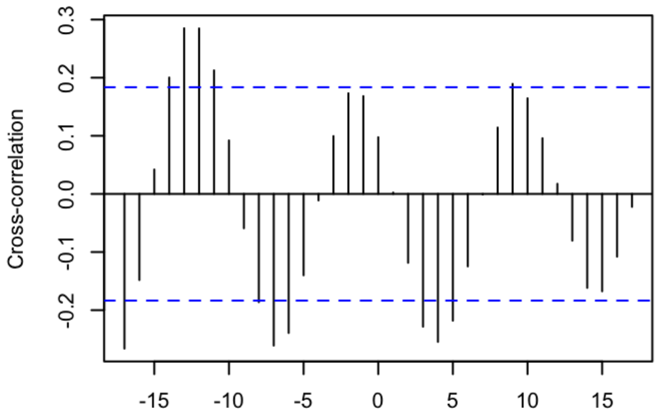
$$r_k^{x,y} = \frac{g_k^{x,y}}{s_x s_y}$$

Just as with other measures of correlation

$$-1 \leq r_k^{x,y} \leq 1$$

# Example of cross-correlation



# SOME SIMPLE MODELS

# White noise (WN)

A time series $\{w_t\}$ is discrete white noise if its values are

1. independent
2. identically distributed with a mean of zero

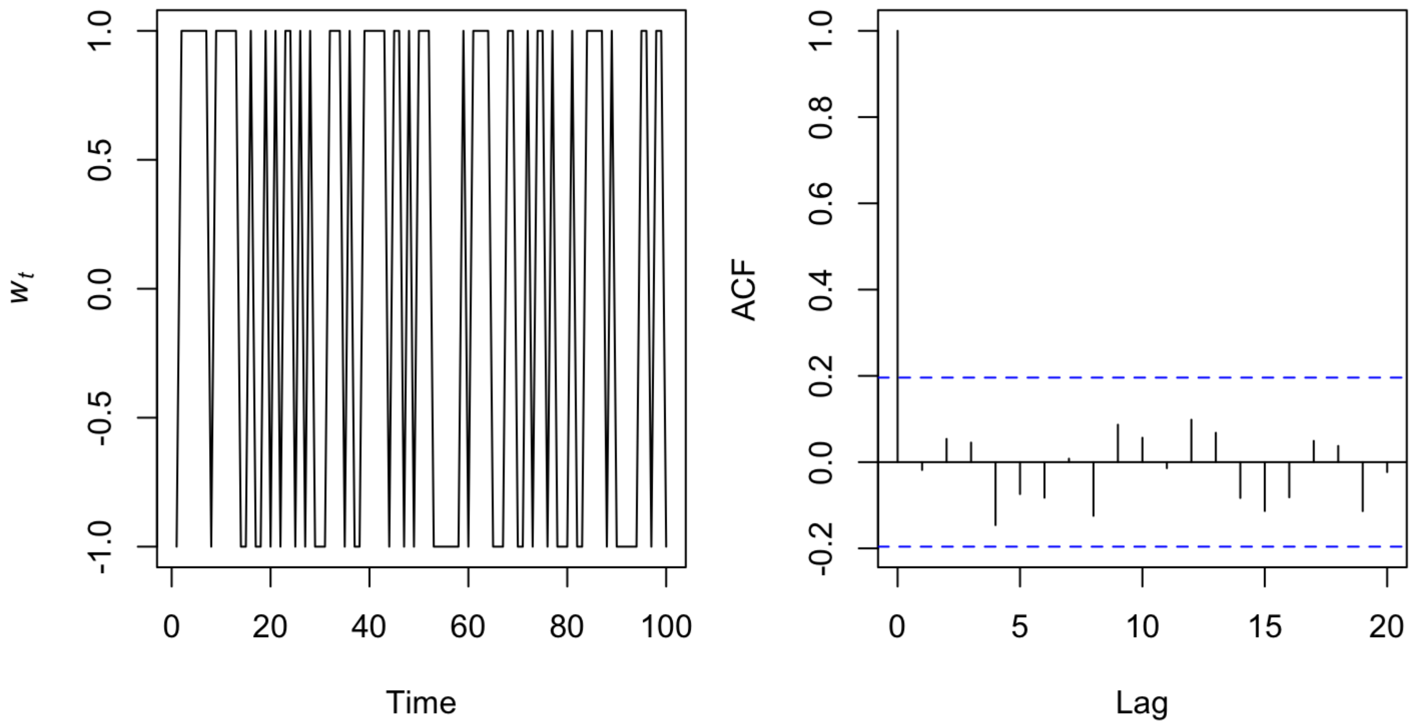
# White noise (WN)

A time series $\{w_t\}$ is discrete white noise if its values are

1. independent
2. identically distributed with a mean of zero

Note that distributional form for $\{w_t\}$ is flexible

# White noise (WN)




$w_t = 2e_t - 1; e_t \sim \text{Bernoulli}(0.5)$

# Gaussian white noise

We often assume so-called *Gaussian white noise*, whereby

$$w_t \sim \mathrm{N}(0, \sigma^2)$$

# Gaussian white noise

We often assume so-called *Gaussian white noise*, whereby

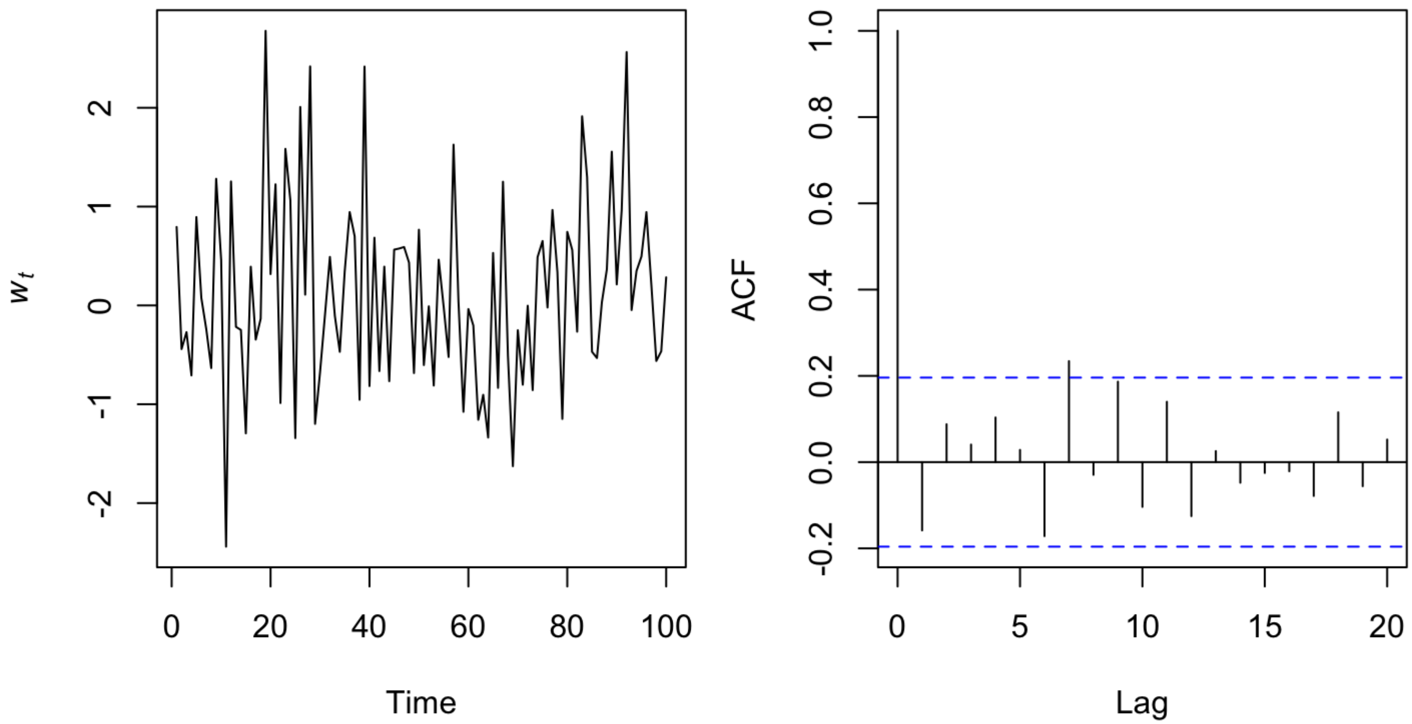
$$w_t \sim \text{N}(0, \sigma^2)$$

and the following apply as well

autocovariance: $\gamma_k = \begin{cases} \sigma^2 & \text{if } k = 0 \\ 0 & \text{if } k \geq 1 \end{cases}$

autocorrelation: $\rho_k = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{if } k \geq 1 \end{cases}$

# Gaussian white noise




$w_t \sim \mathrm{N}(0, 1)$

# Random walk (RW)

A time series $\{x_t\}$ is a random walk if

1. $x_t = x_{t-1} + w_t$
2. $w_t$ is white noise

# Random walk (RW)

The following apply to random walks

mean: $\mu_x = 0$

autocovariance: $\gamma_k(t) = t\sigma^2$

autocorrelation: $\rho_k(t) = \frac{t\sigma^2}{\sqrt{t\sigma^2(t+k)\sigma^2}}$

# Random walk (RW)

The following apply to random walks

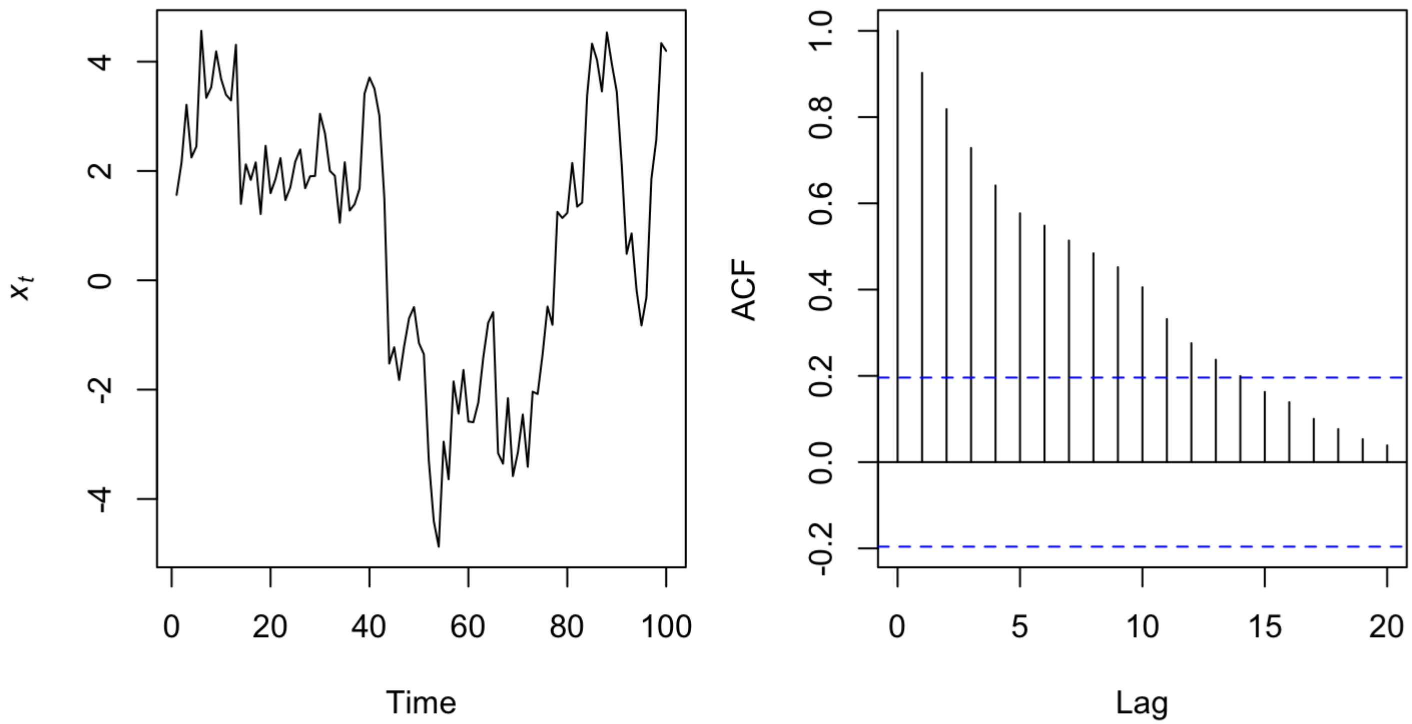
mean: $\mu_x = 0$

autocovariance: $\gamma_k(t) = t\sigma^2$

autocorrelation: $\rho_k(t) = \frac{t\sigma^2}{\sqrt{t\sigma^2(t+k)\sigma^2}}$

*Note*: Random walks are not stationary

# Random walk (RW)




$$x_t = x_{t-1} + w_t; w_t \sim \mathrm{N}(0, 1)$$

# SOME IMPORTANT OPERATORS

# The backshift shift operator

The *backshift shift operator* ($\mathbf{B}$) is an important function in time series analysis, which we define as

$$\mathbf{B}x_t = x_{t-1}$$

or more generally as

$$\mathbf{B}^k x_t = x_{t-k}$$

# The backshift shift operator

For example, a random walk with

$$x_t = x_{t-1} + w_t$$

can be written as

$$\begin{aligned} x_t &= \mathbf{B}x_t + w_t \\ x_t - \mathbf{B}x_t &= w_t \\ (1 - \mathbf{B})x_t &= w_t \\ x_t &= (1 - \mathbf{B})^{-1} w_t \end{aligned}$$

# The difference operator

The *difference operator* ($\nabla$) is another important function in time series analysis, which we define as

$$\nabla x_t = x_t - x_{t-1}$$

# The difference operator

The *difference operator* ($\nabla$) is another important function in time series analysis, which we define as

$$\nabla x_t = x_t - x_{t-1}$$

For example, first-differencing a random walk yields white noise

$$\begin{aligned} \nabla x_t &= x_{t-1} + w_t \\ x_t - x_{t-1} &= x_{t-1} + w_t - x_{t-1} \\ x_t - x_{t-1} &= w_t \end{aligned}$$

# The difference operator

The difference operator and the backshift operator are related

$$\nabla^k = (1 - \mathbf{B})^k$$

# The difference operator

The difference operator and the backshift operator are related

$$\nabla^k = (1 - \mathbf{B})^k$$

For example

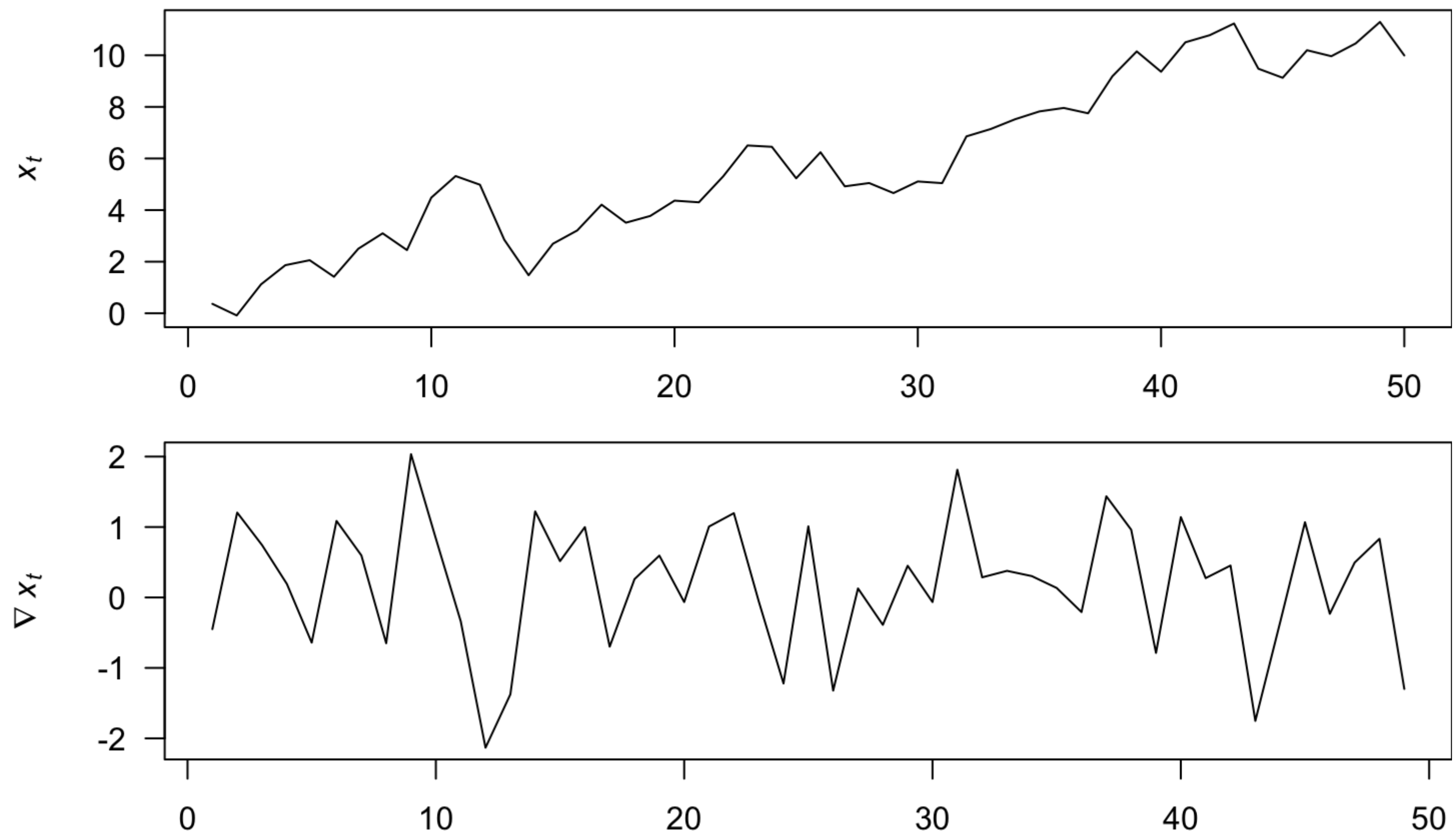
$$\begin{aligned} \nabla x_t &= (1 - \mathbf{B}) x_t \\ x_t - x_{t-1} &= x_t - \mathbf{B} x_t \\ x_t - x_{t-1} &= x_t - x_{t-1} \end{aligned}$$

# Differencing to remove a trend

Differencing is a simple means for removing a trend

The 1st-difference removes a linear trend; a 2nd-difference would remove a quadratic trend, etc.

# Differencing to remove a trend

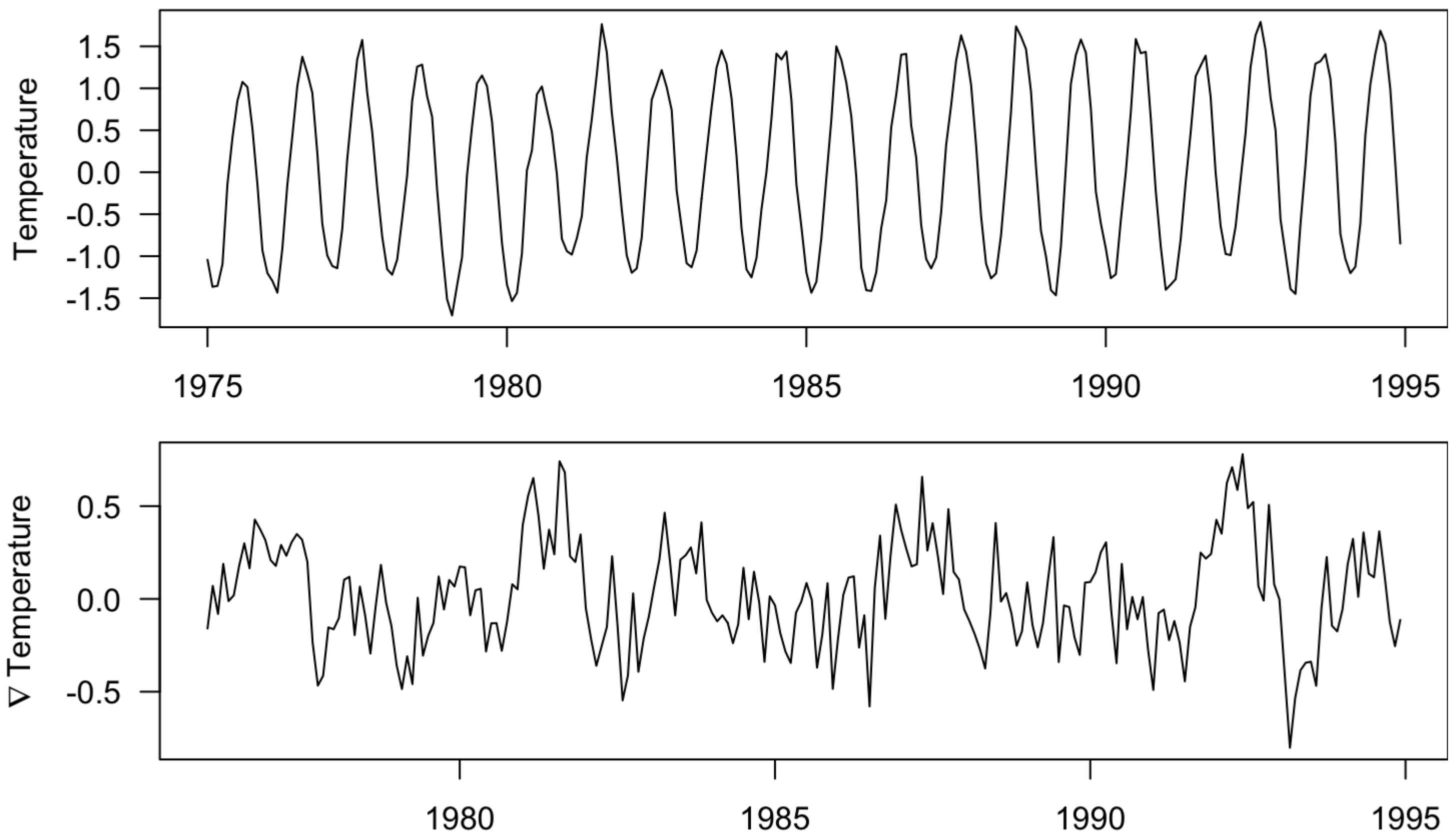


# Differencing to remove seasonality

Differencing is a simple means for removing a seasonal effect

Using a 1st-difference with $k = period$ removes both trend & seasonal effects

# Differencing to remove seasonality



# Topics for today

Characteristics of time series

- Expectation, mean & variance
- Covariance & correlation
- Stationarity
- Autocovariance & autocorrelation
- Correlograms

White noise

Random walks

Backshift & difference operators