Time Series Analysis

Lecture 1: Introduction

Tohid Ardeshiri

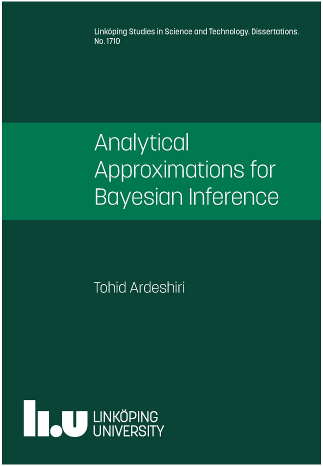
Linköping University
Division of Statistics and Machine Learning

September 2, 2019

Course teacher

Tohid Ardeshiri
PhD in 2015 in Bayesian inference

Senior Data Scientist at
Qamcom Research & Technology AB



Bayesian Inference

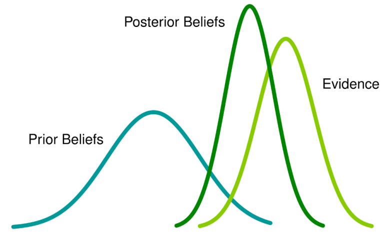
Bayesian inference is a means of combining prior beliefs with the data (evidence) to obtain posterior beliefs.

Example: Parameter learning

$$f(\boldsymbol{\theta}|\mathbf{x}) \propto f(\mathbf{x}|\boldsymbol{\theta})f(\boldsymbol{\theta})$$

Probability Calculus

$$f(\boldsymbol{\theta},\mathbf{x}) = f(\mathbf{x}|\boldsymbol{\theta})f(\boldsymbol{\theta})$$
$$f(\boldsymbol{\theta},\mathbf{x}) = f(\boldsymbol{\theta}|\mathbf{x})f(\mathbf{x})$$

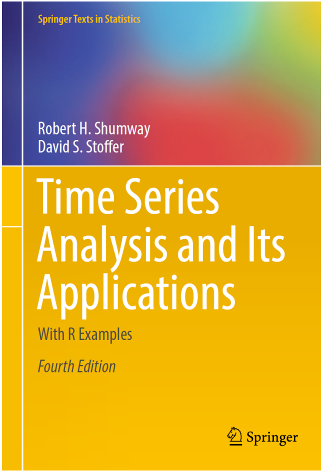


Course literature and software

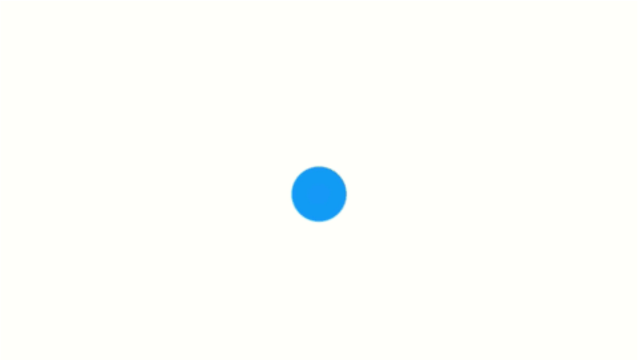
Course literature:
Time series Analysis and its Applications
Can be downloaded freely here:

https://www.stat.pitt.edu/stoffer/tsa4/tsa4.pdf

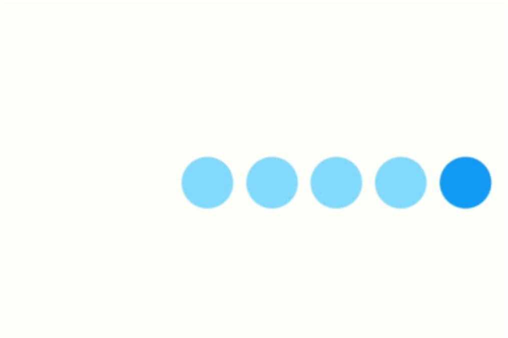
Software for computer labs is **R**:



Sequential data



Sequential data: Motion of a ball



Sequential data: A sentence

This is a sequential data type.

Sequential data: A sentence

This is a sequential data type.

This	is	a	sequential	data	type	.

Sequential data: A sentence or a word

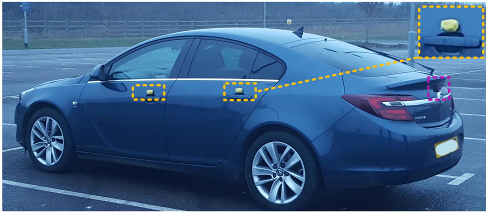
This is a sequential data type.

This | is | a | sequential | data | type | .

s | e | q | u | e | n | t | i | a | l

A look at real data

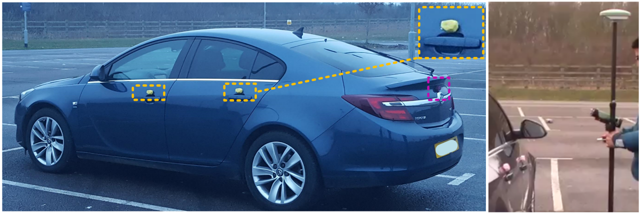
Received signal strength indicator (RSSI) is a common observation (data).



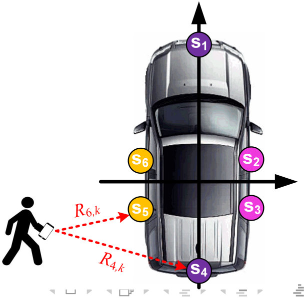
Where is the driver?

A look at real data

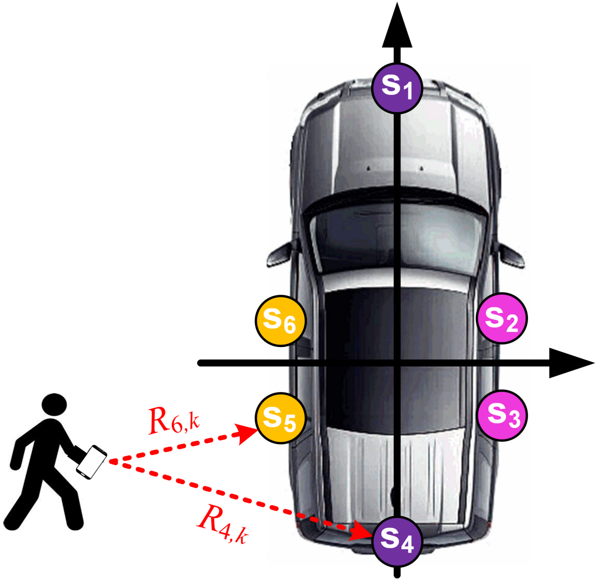
Received signal strength indicator (RSSI) is a common observation (data).



Where is the driver?



Where is the driver?

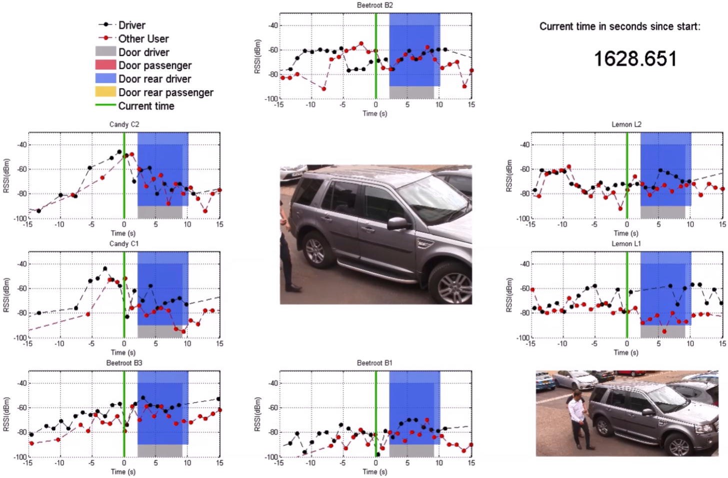


Where is the driver?

Video of data collection

Where is the driver?

Animation of the of signals

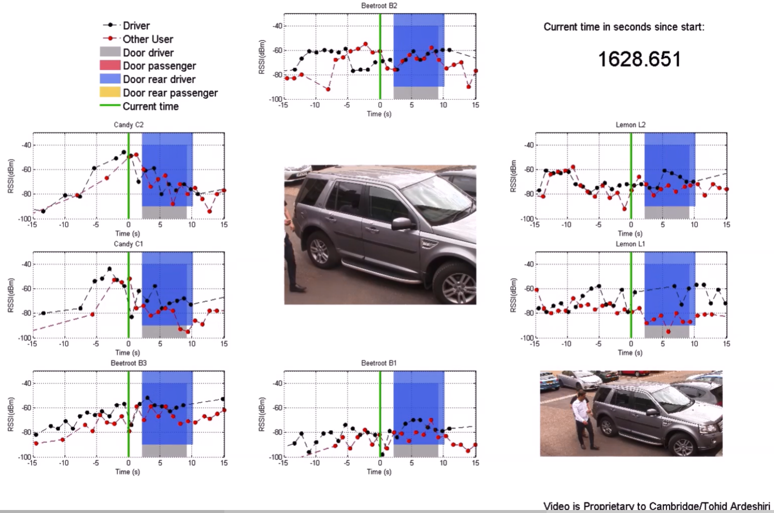


Video is Proprietary to Cambridge/Tohid Ardeshiri

Time Series Analysis

What is a Time Series?

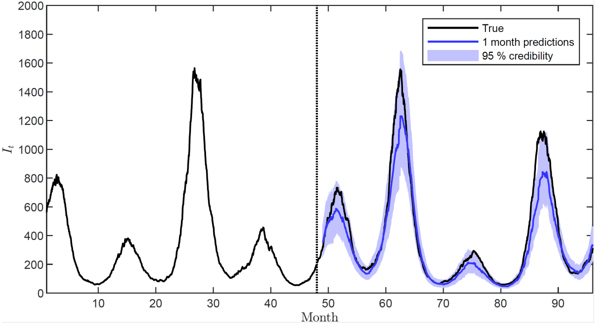
- A sequential data where observations are collected over time
- Observations are typically correlated!



Time Series Analysis

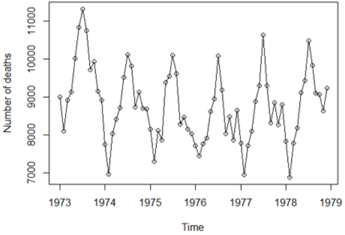
- Understand the properties of the underlying process
- Be able to predict (forecast) possible future values
- Reason about the **uncertainties** in the predictions

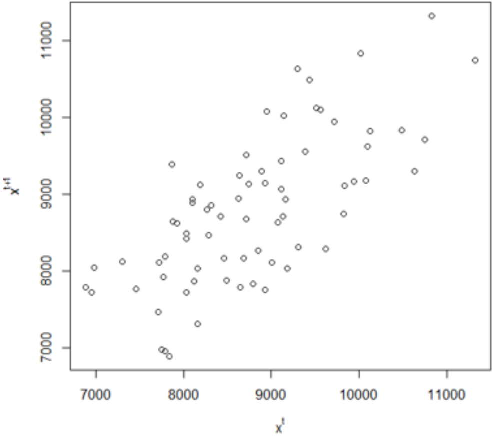
  requires statistical methods!



Time Series Analysis

Usual regression analysis: observations are often **iid.**
Time Series Analysis: observations are **correlated!**

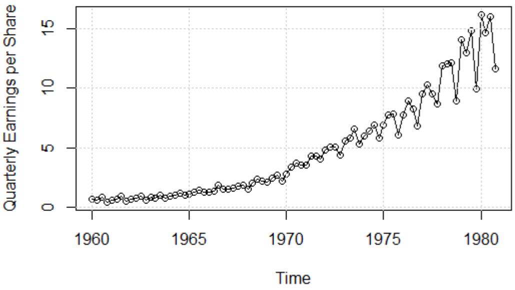




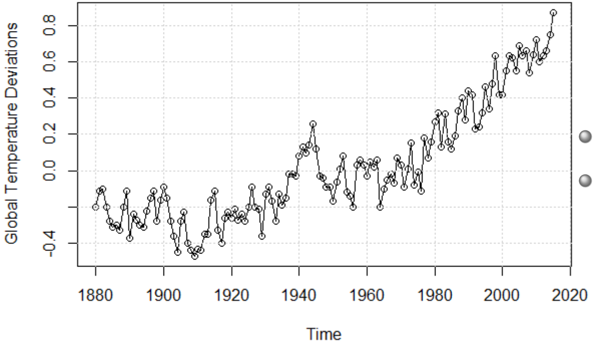
Ex) See connection between x_t and x_{t+1}

Ex 1: Johnson & Johnson quarterly earnings

- **What can you see?**
  - Trend?
    - Constant
    - Linear
    - Other
  - Variation?
  - Seasonality?
  - Outliers?

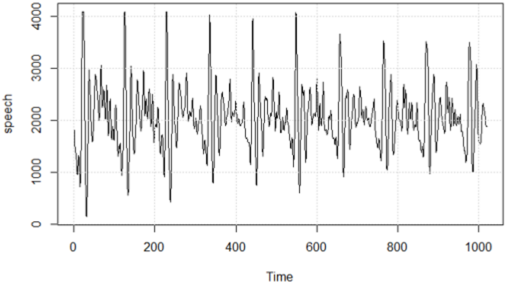


Ex 2: Global warming



- **What can you see?**
- Trend is of interest

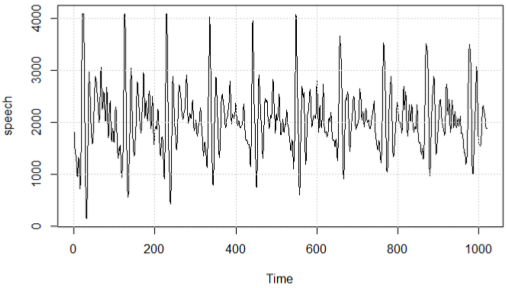
Ex 3: Speech data



- **What can you see?**

Pattern of periodicity is of interest $\rightarrow$ decompose signal into different frequencies

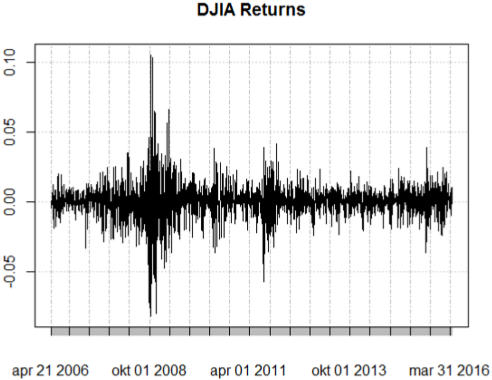
Ex 3: Speech data



- **What can you see?**

Pattern of periodicity is of interest → decompose signal into different frequencies
not covered in this course!

Ex 4: Dow Jones Industrial Average

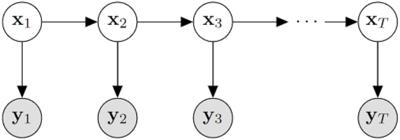


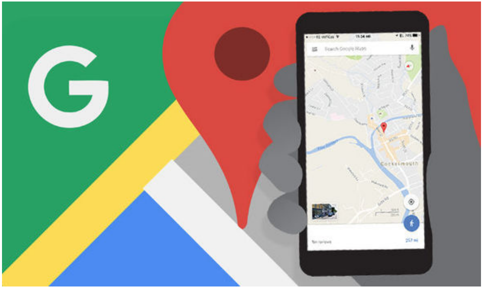
- **What can you see here?**

Pattern of periodicity is of interest → Stochastic volatility

Ex 5: Dynamical systems

Linear and Guassian state-space models for tracking objects

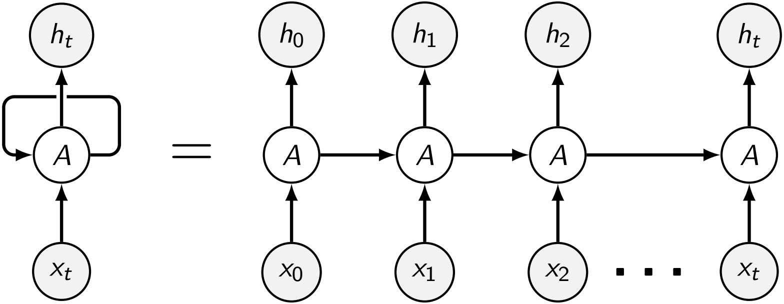




Ex 6: Recurrent neural networks

Natural Language Processing

This is a sequential data type .



Time Series Analysis

Application areas

- Natural sciences
- Climatology
- Robotics/autonomous systems
- Social sciences
- Medicine
- Economics
- Telecommunications
- ...

The Big Picture

Time Series Analysis

Time Domain
- ARIMA models
- State-space models
- (RNNs)

Frequency domain
- Spectral estimation
- Frequency domain filtering

This course

Time domain: The Big Picture

Time Series data

Exploratory data analysis

Make data stationary

Suggest a model $M(\phi, \theta)$

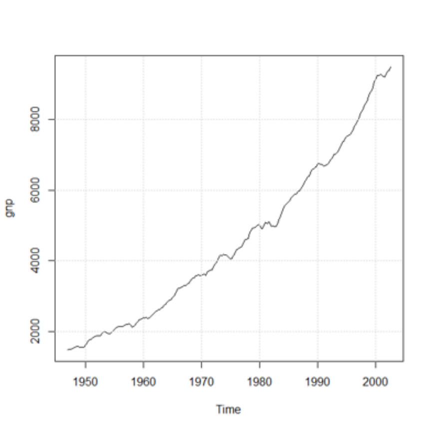
$\phi = (\phi_1, ..., \phi_p)$

$\theta = (\theta_1, ..., \theta_q)$

Prediction

Mod

- A
  ar



- Estimate $\phi' s$, $\theta' s$, ...

Time domain: The Big Picture

Time Series data

Exploratory data analysis

Make data stationary

Suggest a model M(ϕ, θ)

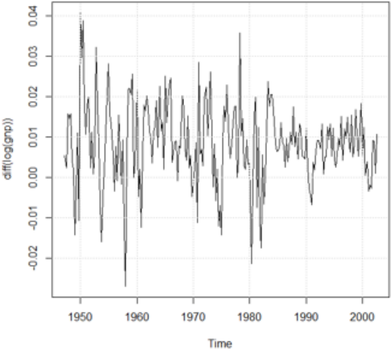
$\phi = (\phi_1, ..., \phi_p)$

$\theta = (\theta_1, ..., \theta_q)$

Estimate $\phi's$, $\theta's$, ...

Prediction

$$Y_t = \nabla(\log(X_t))$$



Time domain: The Big Picture

Time Series data

Exploratory data analysis

Make data stationary

Suggest a model $M(\phi, \theta)$

$$\phi = (\phi_1, ..., \phi_p)$$
$$\theta = (\theta_1, ..., \theta_q)$$

Model Estimation

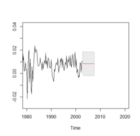
- Estimate $\phi's$, $\theta's$, ...

Model diagnostics and Model selection

- Are there better models with other p and q or even different kind?

no

Prediction



Time domain: The Big Picture

Time Series data

Exploratory data analysis

Make data stationary

Suggest a model $\mathrm{M}(\phi, \theta)$

$$\phi = (\phi_1, ..., \phi_p)$$

$$\theta = (\theta_1, ..., \theta_q)$$

Model Estimation

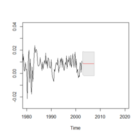
- Estimate $\phi's$, $\theta's$, ...

Model diagnostics and Model selection

- Are there better models with other p and q or even different kind?

no

Prediction



Time domain: The Big Picture

Time Series data

Exploratory data analysis

Make data stationary

Suggest a model $\mathrm{M}(\phi, \theta)$

$$\phi = (\phi_1, ..., \phi_p)$$
$$\theta = (\theta_1, ..., \theta_q)$$

Model Estimation

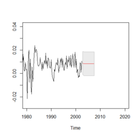
- Estimate $\phi's$, $\theta's$, ...

Model diagnostics and Model selection

- Are there better models with other p and q or even different kind?

no

Prediction



Time domain: The Big Picture

Time Series data

Exploratory data analysis

Make data stationary

Suggest a model M(ϕ, θ)

$$\phi = (\phi_1, ..., \phi_p)$$

$$\theta = (\theta_1, ..., \theta_q)$$

Model Estimation

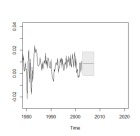
- Estimate $\phi's$, $\theta's$, ...

Model diagnostics and Model selection

- Are there better models with other p and q or even different kind?

no

Prediction



Time domain: The Big Picture

Time Series data

Exploratory data analysis

Make data stationary

Suggest a model $M(\phi, \theta)$

$$\phi = (\phi_1, ..., \phi_p)$$
$$\theta = (\theta_1, ..., \theta_q)$$

Model Estimation

- Estimate $\phi's$, $\theta's$, ...

Model diagnostics and Model selection

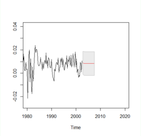
- Are there better models with other p and q or even different kind?

yes

Revise the model

no

Prediction



Course topics

- Time series regression and explorative analysis
- ARIMA models
  - AR, MA, ARMA, ARIMA, seasonal ARIMA
  - Model selection
  - Estimation
  - Forecasting
- State space models
  - Linear and Gaussian state space models
  - Kalman filtering and smoothing
- Recurrent Neural Networks (RNNs)

Course organization

- Lectures
  - Available at LISAM
- Teaching sessions
- Computer labs
  - Available at LISAM, under Submissions
  - Work in pairs
  - Send your report via LISAM
  - Deadlines
- Written assignments
  - Submissions needed - keys are given for some assignments
- Examination
  - Computer based exam
  - Submission of lab reports and written assignments

Course organization

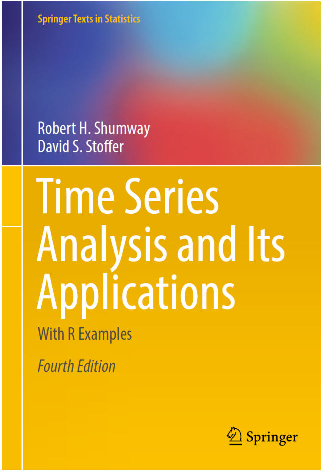
- Software: R
  - https://www.r-project.org/
  - https://www.rstudio.com/
- Define your groups (2 persons) this week:
  - https://docs.google.com/spreadsheets/d/1tzG35WSDWRhHWFA0cN0L1WUzoYqdn0GhZUwvXz3HhII/edit?usp=sharing
  - **Difficult to find a group? Put your name in some cell.. I will merge you to someone**

Course organization

Course literature:
Time series Analysis and its Applications, Fourth Edition (2017), ISBN 978-3-319-52451-1
Can be downloaded freely here:

https://www.stat.pitt.edu/stoffer/tsa4/tsa4.pdf

- Do not skip examples when you read!
- First 2 chapters are easy, but don't relax!



Time Series models

- Time series x_t : random variable
  - A collection of x_t = stochastic process
  - $t = 0, \pm 1, \pm 2, \ldots$
- (probably) Simplest series: white noise
  - w_t uncorrelated (white: all possible periodic oscillations are present at equal strength)

$$w_t \sim wn(0, \sigma_w^2)$$

  - w_t independent and identically distributed (white independent noise)

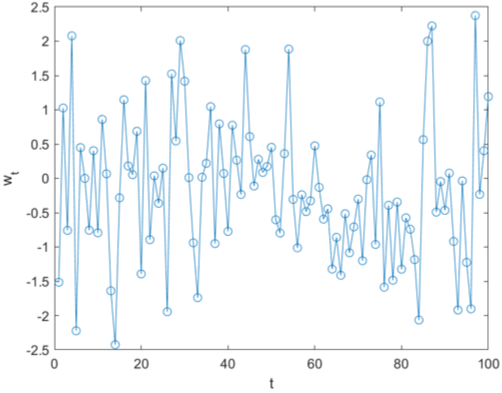
$$w_t \sim iid(0, \sigma_w^2)$$

- Reminder:

$$\text{uncorrelated} \iff E(XY) = EX.EY$$
$$\text{independent} \iff f_{X,Y}(x,y) = f_X(x).f_Y(y)$$

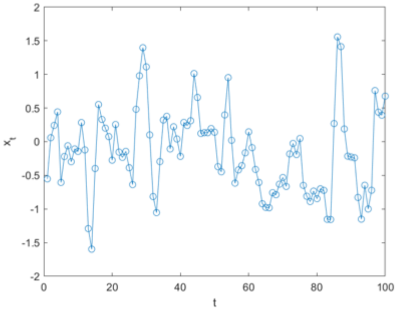
White noise

- Example: $w_t \sim iidN(0, 1)$



Moving average

Example: $x_t = 0.2w_{t-1} + 0.5w_t + 0.2w_{t+1}$



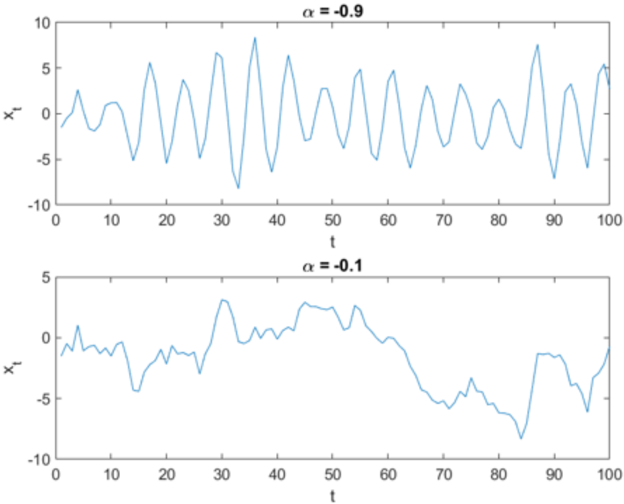
Very Interesting Fact: most stationary processes can be represented as a sum of lagged white noise:

$$x_t = \sum_{j=-\infty}^{\infty} \psi_j w_{t-j}$$

Autoregressive model

Example: AR(2) process (Assume $x_0 = 0, x_{-1} = 0$)

$$x_t = x_{t-1} + \alpha x_{t-2} + w_t$$



Random walk with drift

A simple model for a "drifting" time series

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

- δ is the drift
- $\delta = 0 \Rightarrow$ random walk

Note: if we assume $x_0 = 0$,

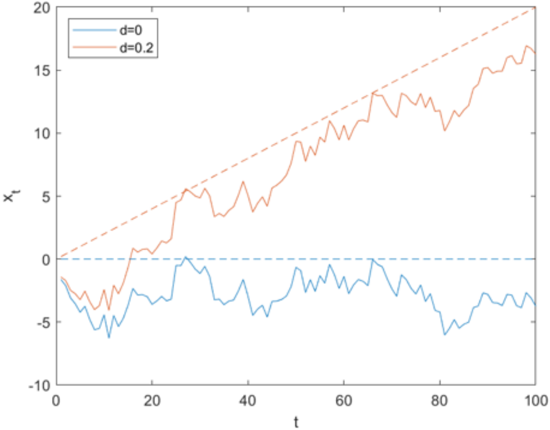
$$x_t = \delta t + \sum_{j=1}^{t} w_j$$

Random walk with drift

A simple model for a "drifting" time series

δ=0 and δ=0.2

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



- δ is the drift
- $\delta = 0 \Rightarrow$ random walk

Note: if we assume $x_0 = 0$,

$$x_t = \delta t + \sum_{j=1}^{t} w_j$$

Basic statistics - reminder

- Probability density function for x: $f(x)$
- Marginal density $f_i(x_i) = \int f(x)dx_1 ... dx_{i-1}dx_{i+1} ... dx_p$
- Expected (mean) value $Ex = \int xf(x)dx$
- Covariance $\operatorname{cov}(x, y) = E\{(x - Ex)(y - Ey)\}$
- Variance $\operatorname{var}(x) = E\{(x - Ex)^2\} \operatorname{cov}(x, x)$
- Relationships (a is a constant)
  - E(x+a)=Ex+a, E(ax)=aEx
  - E(x+y)=Ex+Ey
  - $\operatorname{cov}(x + a, y) = \operatorname{cov}(x, y)$
  - $\operatorname{cov}(x + z, y) = \operatorname{cov}(x, y) + \operatorname{cov}(z, y)$
  - $\operatorname{var}(ax) = a^2 \operatorname{var}(x)$

Statistical representation of a time series

Which measures of dependence exist for time series?

- Theoretical?
- Practical?

Given time series $x_1, ...x_n$ measured at fixed $t_1, ...t_n$

- Joint pdf

$$f_{t_1,...t_n}(x_{t_1}, ..., x_{t_n})$$

- Marginal pdf

$$f_t(x_t)$$

Statistical representation of a time series
on whiteboard

Mean function at time t

$$\mu_t = E(x_t) = \int_{-\infty}^{\infty} x f_t(x) dx$$

Examples: Compute mean function for

- Moving average $x_t = 0.2w_{t-1} + 0.5w_t + 0.2w_{t+1}$
- Random walk $x_t = \delta t + \sum_{j=1}^{t} w_j$

Autocovariance and ACF

How do we measure linear dependence between two variables? $\rightarrow$ Covariance or Correlation

How do we measure linear dependence between two time-lags in a time series? In the same way!

- Autocovariance function

$$\gamma(s,t) = \text{cov}(x_s, x_t) = E[(x_s - \mu_s)(x_t - \mu_t)]$$

  Note $\text{var}(x_t) = \gamma(t,t)$

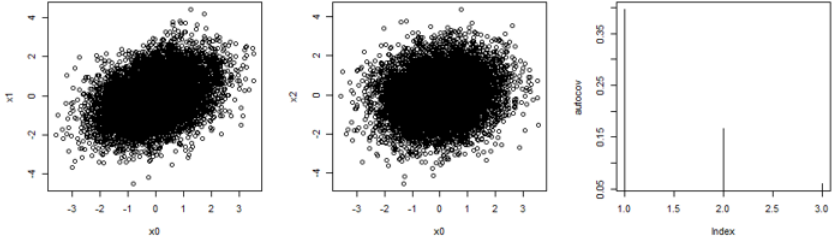
- Autocorrelation function (ACF)

$$\rho(s,t) = \frac{\gamma(s,t)}{\sqrt{\gamma(s,s)\gamma(t,t)}}$$

Autocovariance and ACF

Generate x_0, x_1, x_2 from $x_t = 0.4x_{t-1} + w_t$

- Consider $\gamma(0,1), \gamma(0,2)$



Autocovariance and ACF

Useful fact: If $U = \sum_{j=1}^{m} a_j x_j$ and

$$V = \sum_{k=1}^{r} b_k y_k$$

$$\mathrm{cov}(U, V) = \sum_{j=1}^{m} \sum_{k=1}^{r} a_j b_k \, \mathrm{cov}(x_j, y_k)$$

Examples: Autocovariance and ACF of **on whiteboard**

- White noise
- Random walk $x_t = \delta t + \sum_{j=1}^{t} w_j$
- Moving average $x_t = 0.2 w_{t-1} + 0.5 w_t + 0.2 w_{t+1}$

Home reading

- Shumway and Stoffer, chapters 1.1-1.3
- TS functions in R: ts, plot.ts, acf, ts.intersect, filter, ts.plot