Time Series Analysis

Spring 2024

Peder Bacher and Pernille Yde Nielsen
February 2, 2024

Outline of the lecture

- Welcome to the course
  - Practical information
  - Homepage + Ed discussion
  - Evaluation
- Introduction to time series analysis (Chapter 1)
  - Examples of time series data
- Multivariate random variables (Chapter 2)
  - Probability distributions
  - The multivariate normal distribution
  - Moment representation

Welcome to the course

Peder Bacher

Pernille Yde Nielsen

Thomas Chiras

Einer Ari Árnason

Peter Grønning

Dimitrios Sousounis

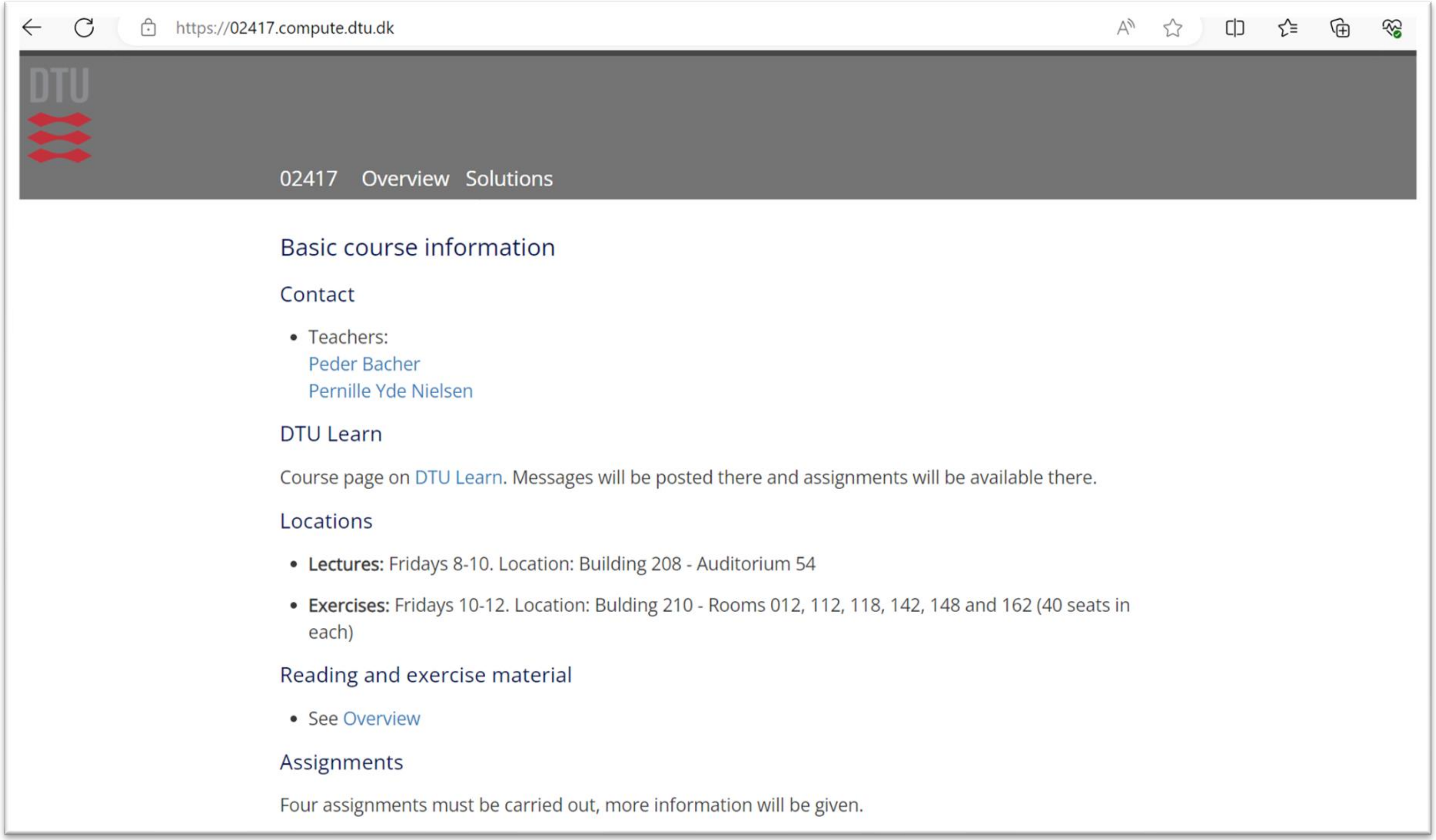
Oscar B.Ø. Pedersen

Hjalte Overgård Pind

Justinas Smertinas

Course homepage

https://02417.compute.dtu.dk/



https://02417.compute.dtu.dk

DTU

02417 Overview Solutions

Basic course information

Contact

- Teachers:
  Peder Bacher
  Pernille Yde Nielsen

DTU Learn

Course page on DTU Learn. Messages will be posted there and assignments will be available there.

Locations

- **Lectures:** Fridays 8-10. Location: Building 208 - Auditorium 54
- **Exercises:** Fridays 10-12. Location: Bulding 210 - Rooms 012, 112, 118, 142, 148 and 162 (40 seats in each)

Reading and exercise material

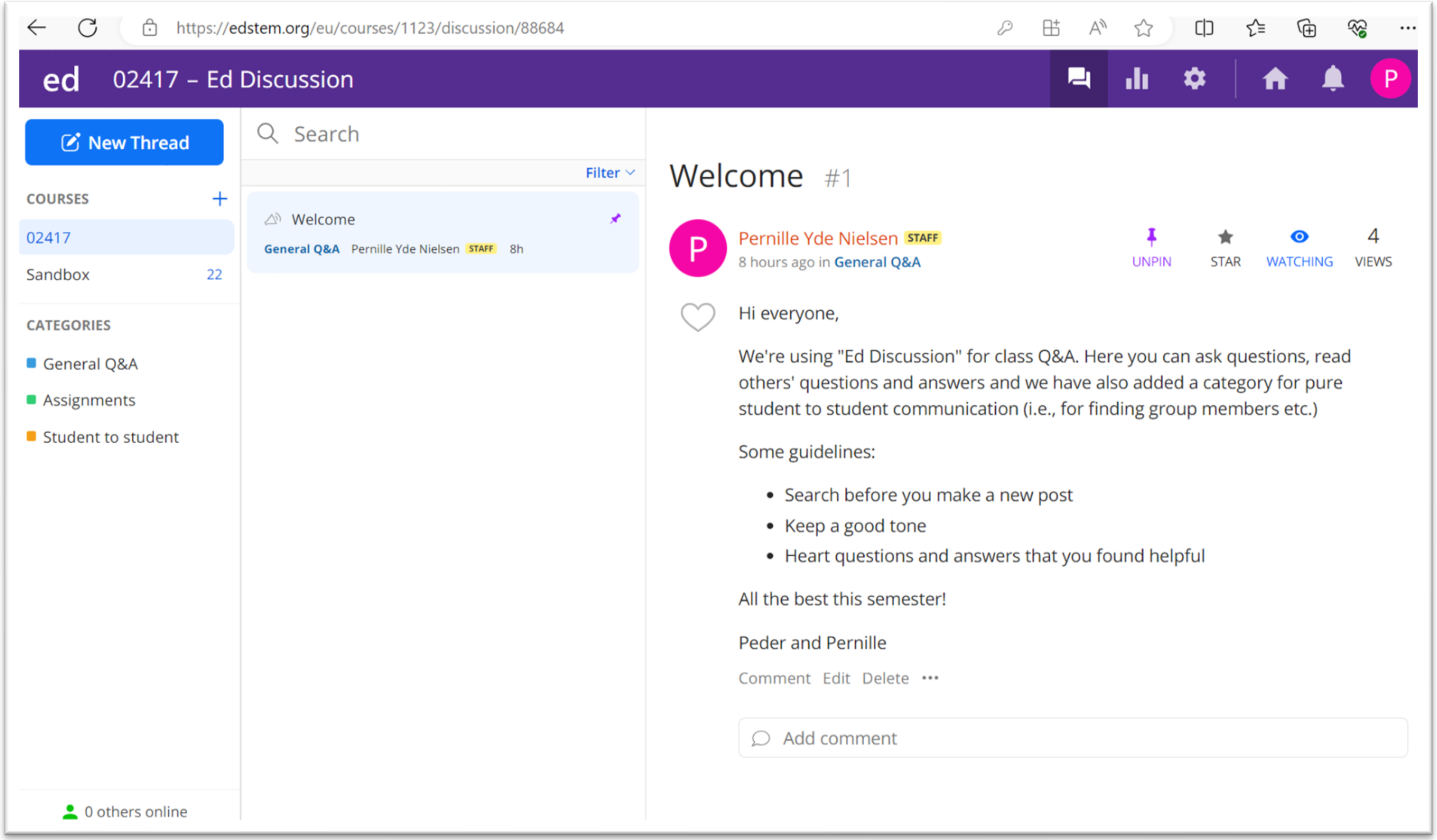
- See Overview

Assignments

Four assignments must be carried out, more information will be given.

Ed Discussion

https://edstem.org/eu/join/BRq6NW



Evaluation

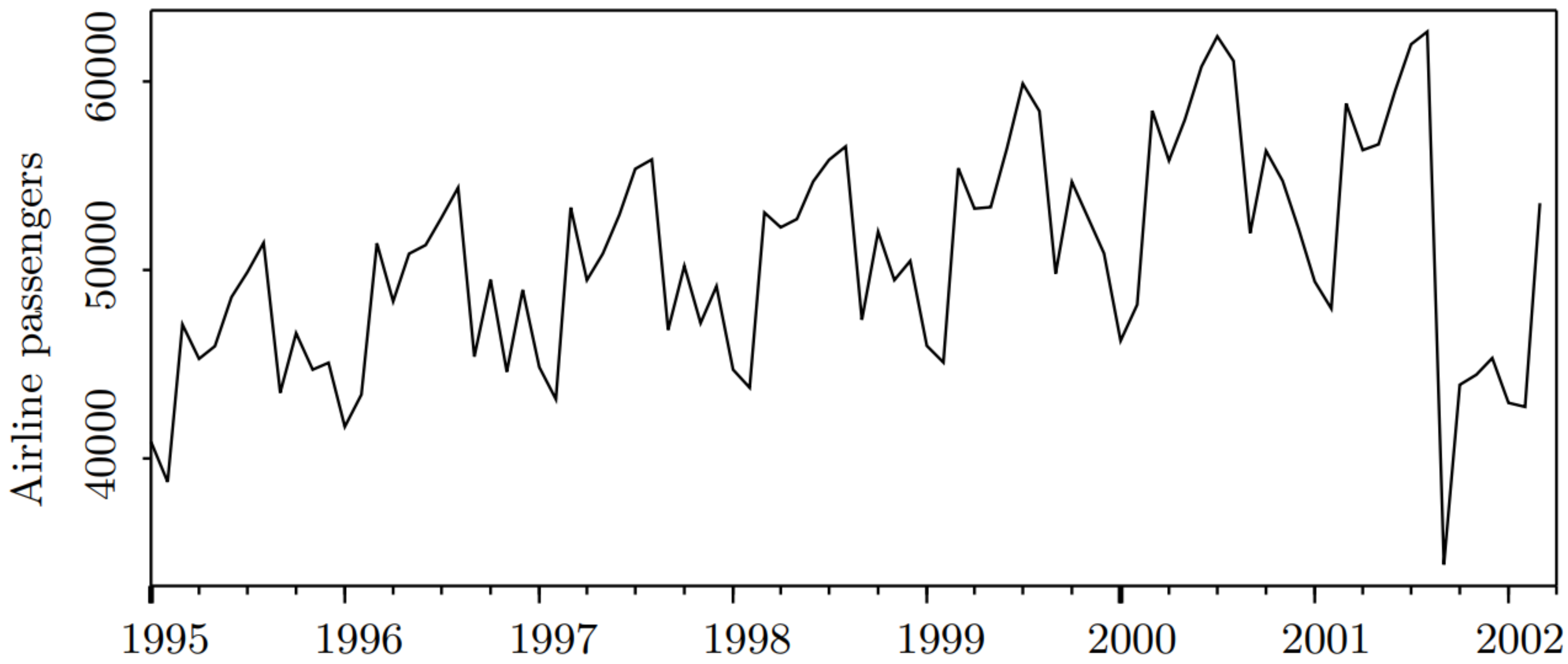
- 4 Assignments
- Peergrade

Introductory example – COLO B shares, 1 month

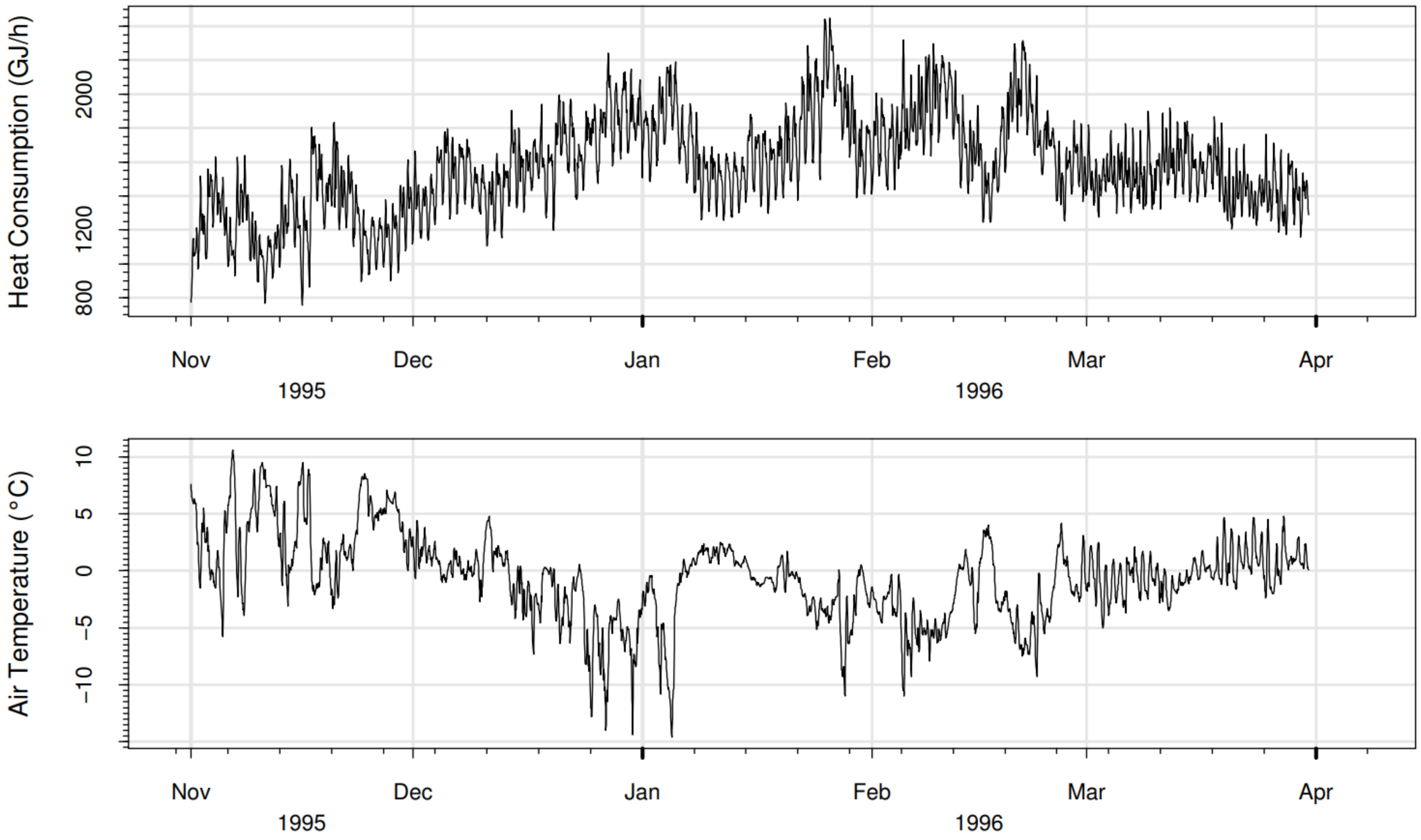
Introductory example – COLO B shares, 1 year

Introductory example – COLO B shares, all

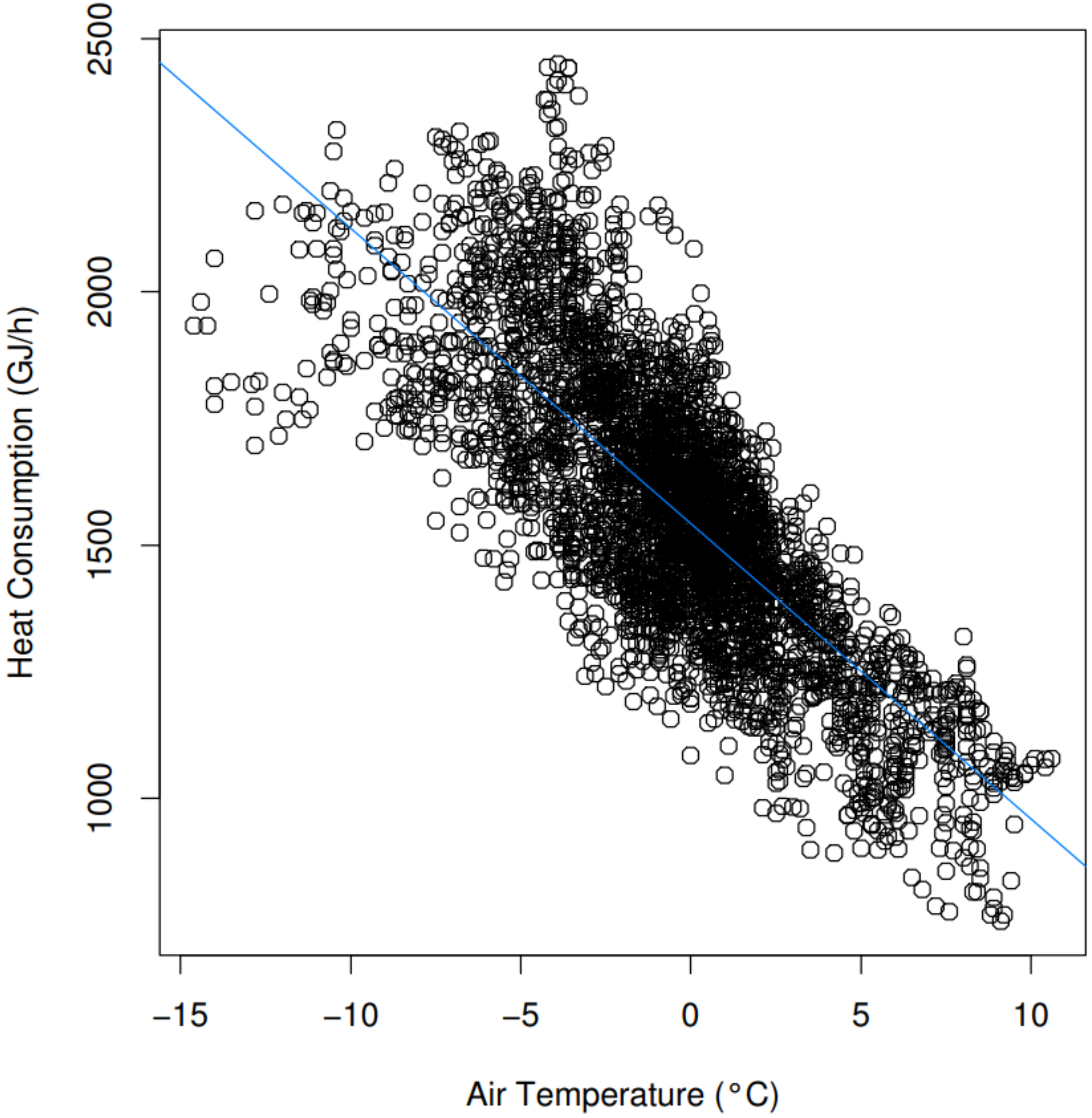
Number of monthly airline passengers in the US



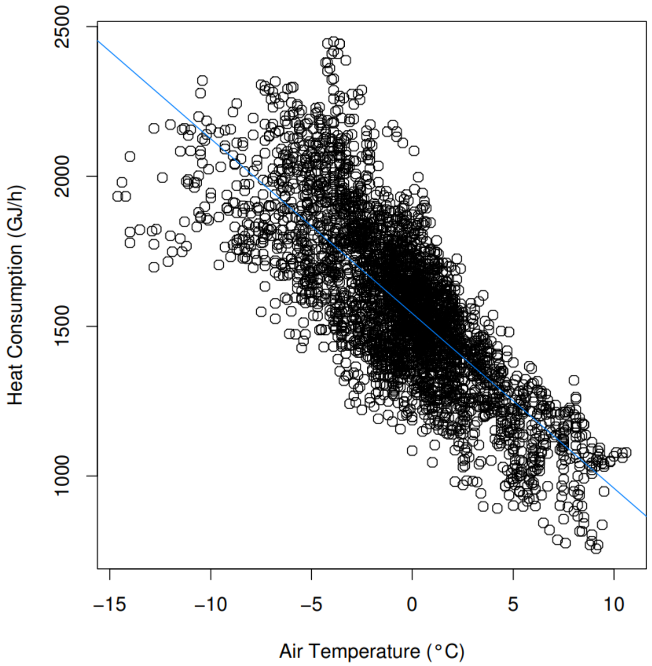
Consumption of district heating (VEKS) data

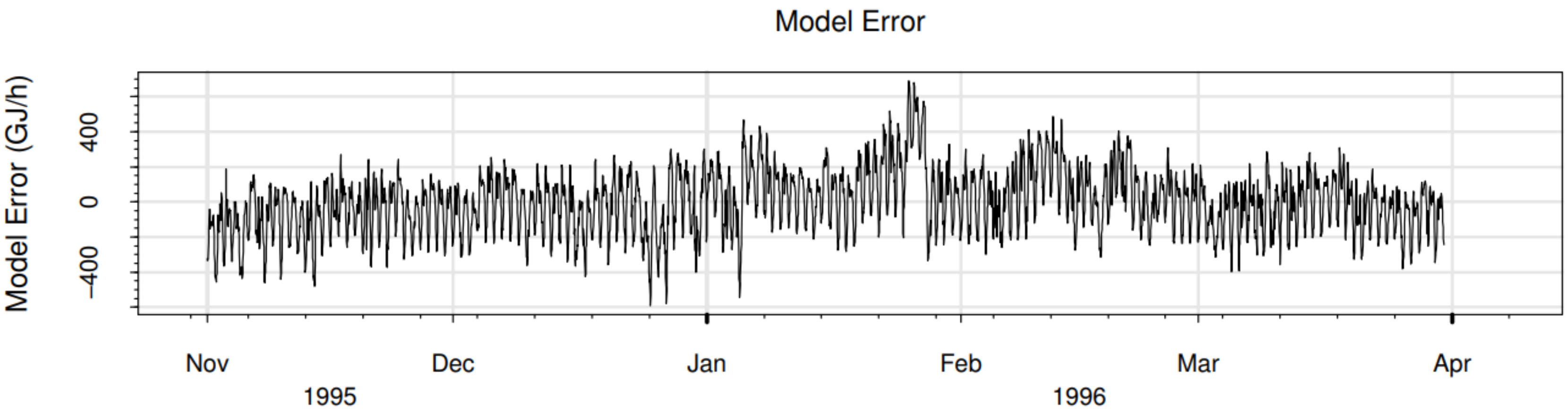


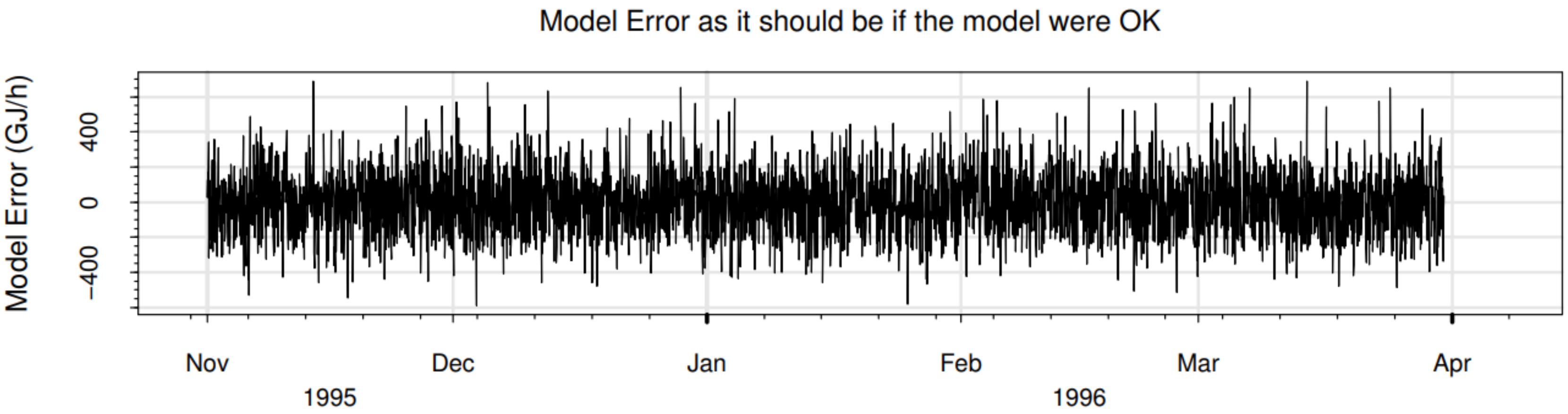
Consumption of district heating – static model



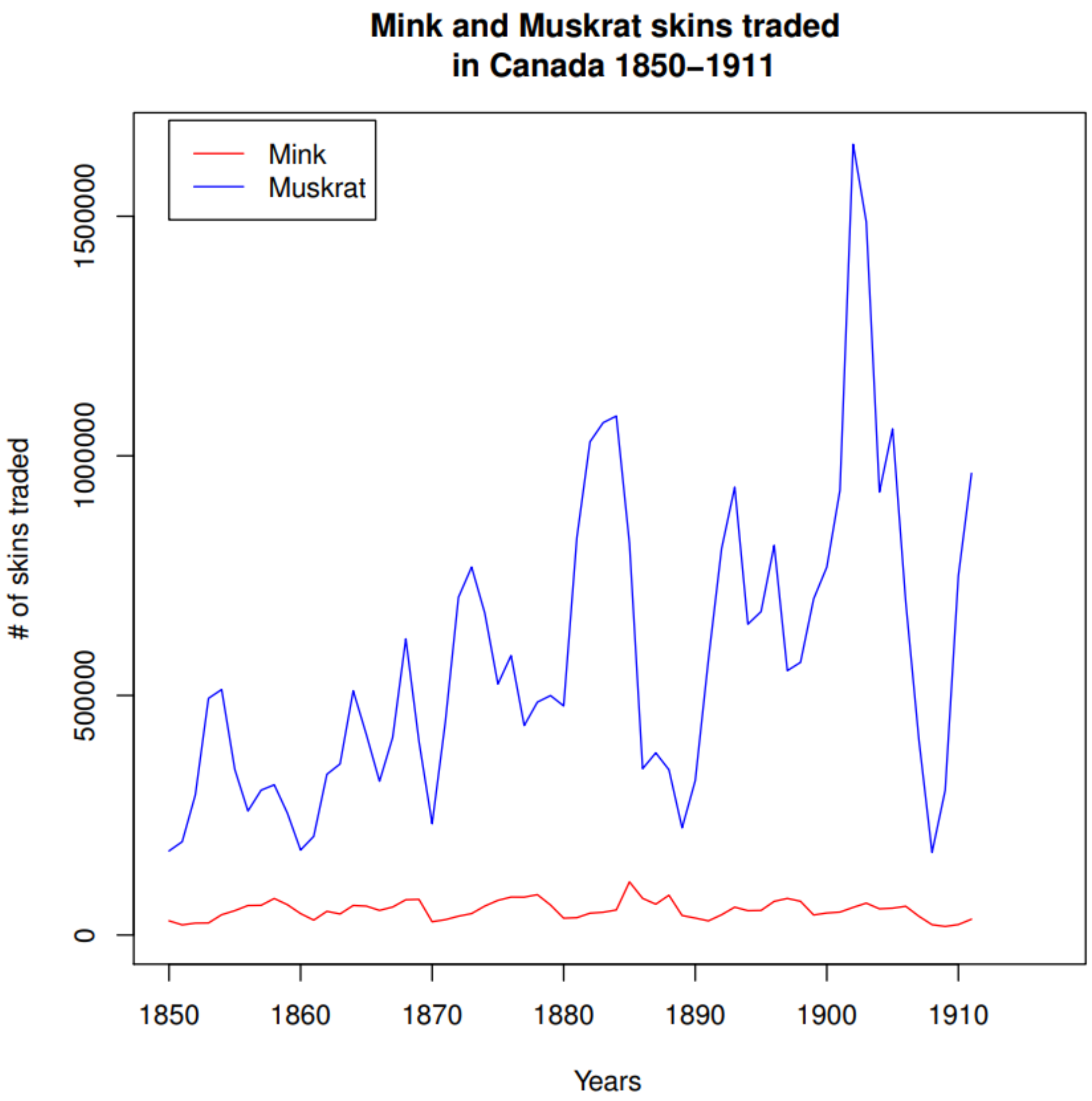
Consumption of district heating – model error



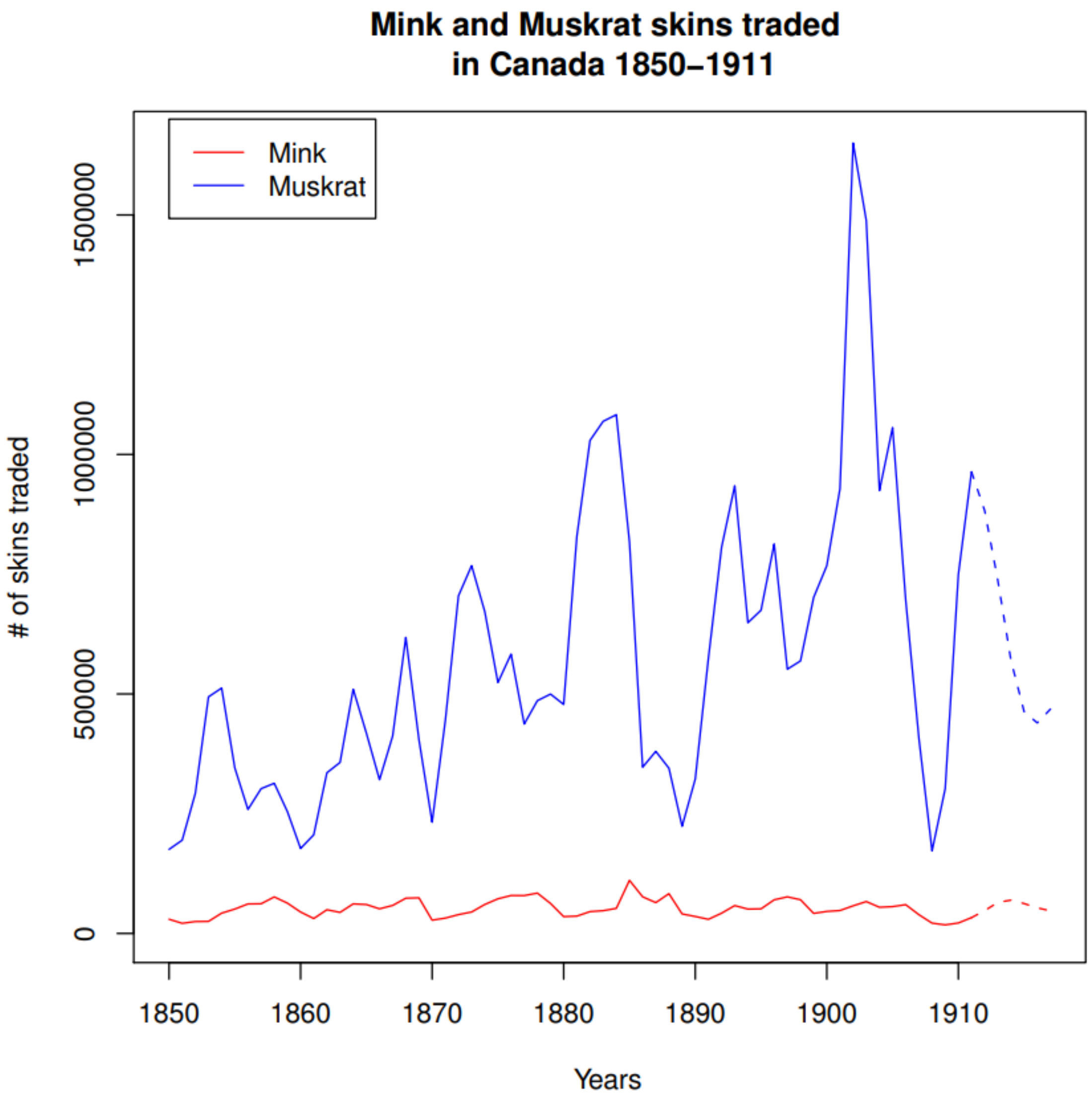




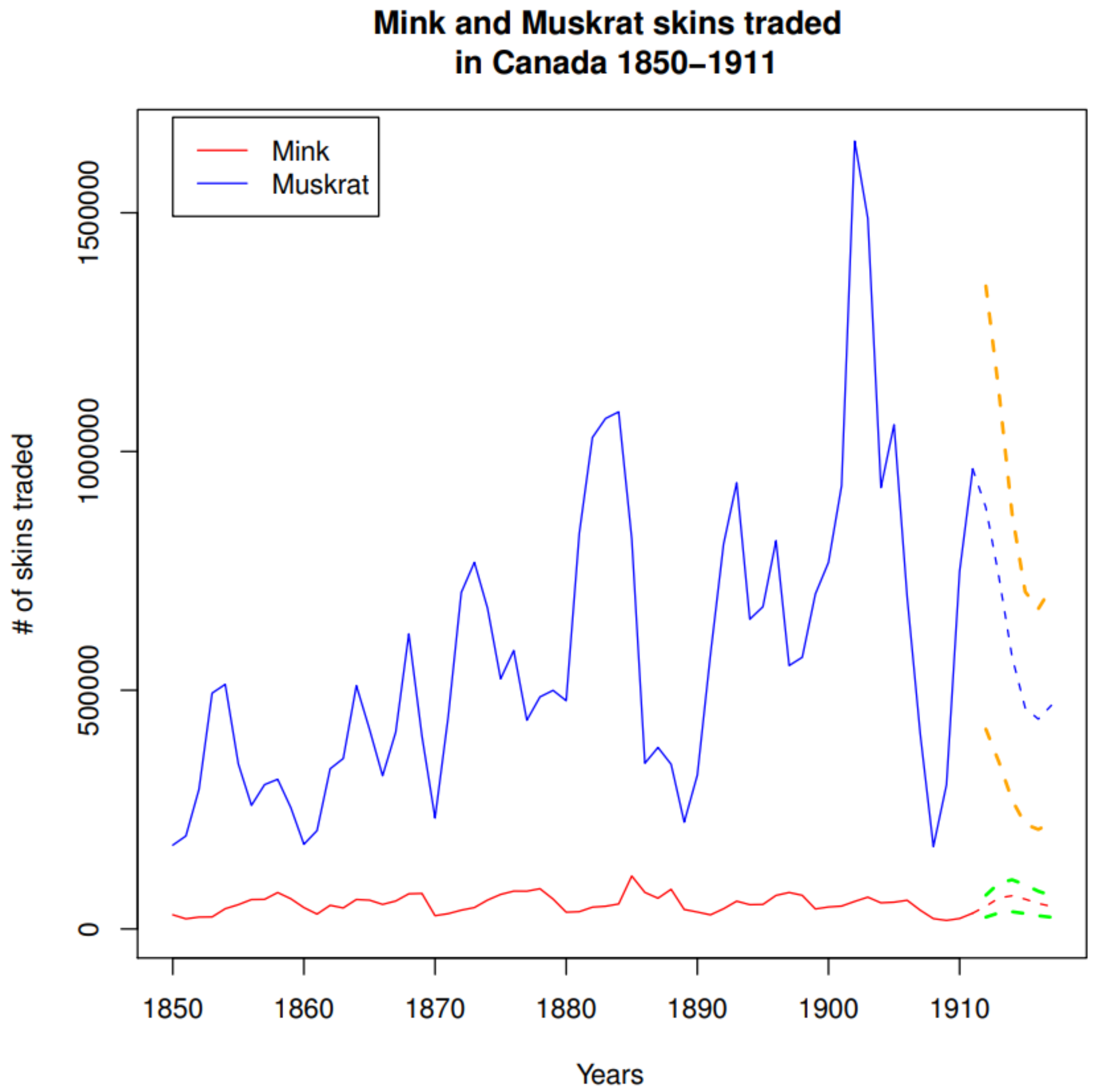
At the end of the course I want you to be able to



Make a forecast



And confidence intervals



Multivariate random variables

- What is a random variable?
- What is a multivariate random variable?
- Why are multivariate random variables essential in time series analysis?

Multivariate random variables

Definition (n-dimensional random variable; random vector)

$$\boldsymbol{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$$

Joint distribution function:

$$F(x_1, \ldots, x_n) = \mathrm{P}\{X_1 \leq x_1, \ldots, X_n \leq x_n\}$$

When does this simplify to

$$F(x_1, \ldots, x_n) = \mathrm{P}\{X_1 \leq x_1\}\mathrm{P}\{X_2 \leq x_2\} \ldots \mathrm{P}\{X_n \leq x_n\}?$$

Joint distribution function and joint density function

Joint distribution function (repeated from last slide):

$$F(x_1, \ldots, x_n) = \mathrm{P}\{X_1 \leq x_1, \ldots, X_n \leq x_n\}$$

Joint density function (often abbreviated **pdf**):

$$f(x_1, \ldots, x_n) = \frac{\partial^n F(x_1, \ldots, x_n)}{\partial x_1 \ldots \partial x_n}$$

(and back to the distribution function:)

$$F(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_n} f(t_1, \ldots, t_n)\, dt_1 \ldots dt_n$$

The multivariate normal distribution

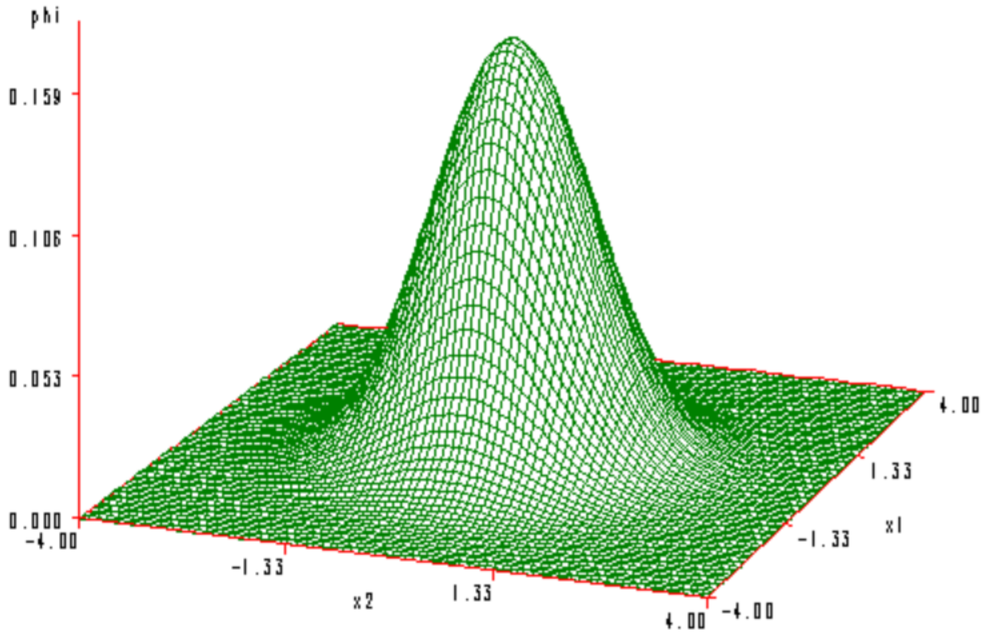
Joint pdf (probability density function):

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}\sqrt{\det \boldsymbol{\Sigma}}} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]$$

$\boldsymbol{\Sigma}$ is the **variance-covariance matrix** (symmetric)

If the variables are un-correlated then the variance-covariance matrix will be a diagonal matrix

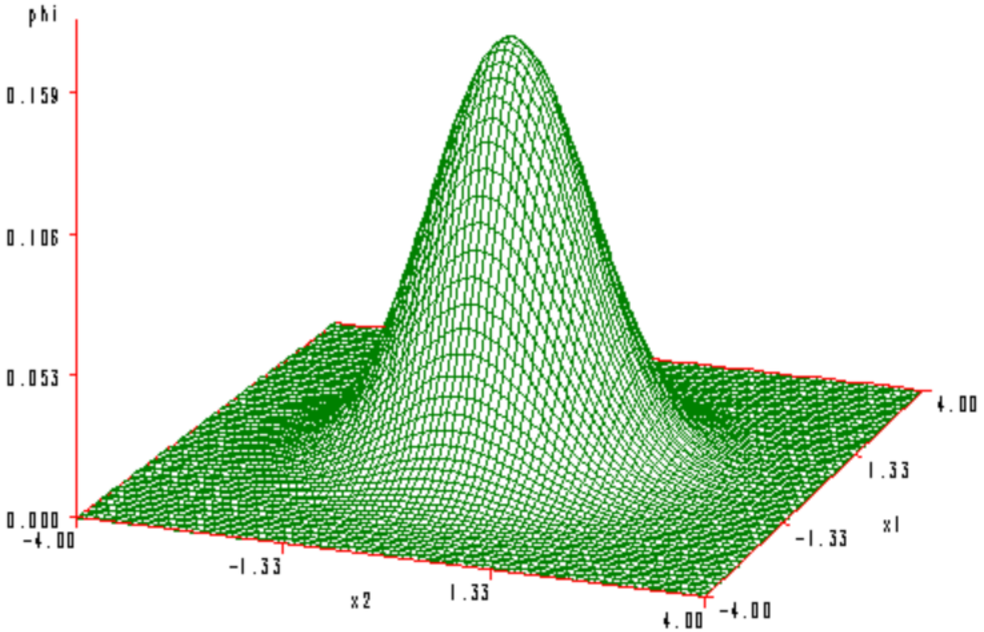
$$\Sigma = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_p^2 \end{pmatrix}$$

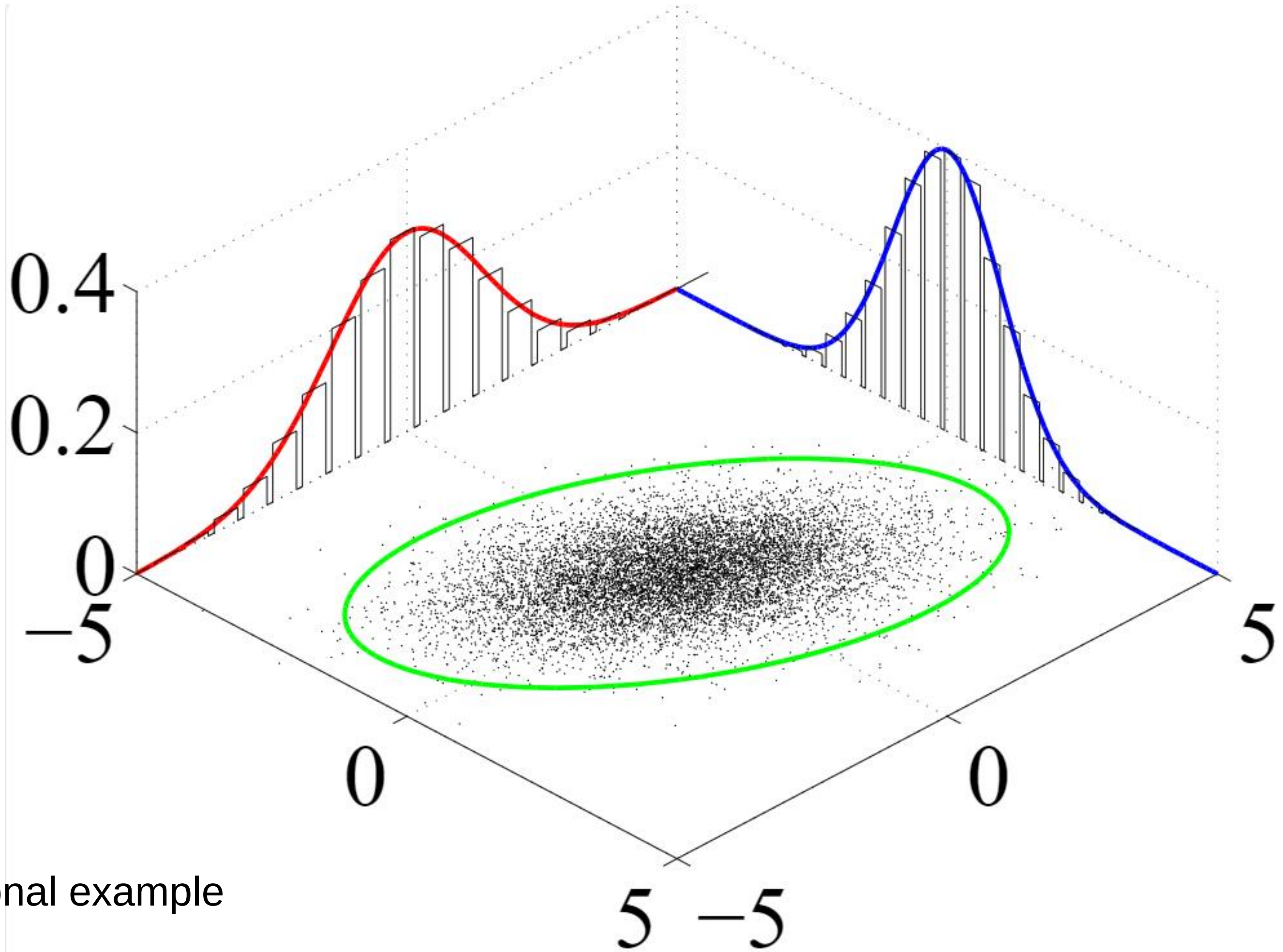


2-dimensional example

Notation: $\boldsymbol{X} \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ means the multivariate random variable **X** is normally distributed

Joint and marginal density function





2-dimensional example

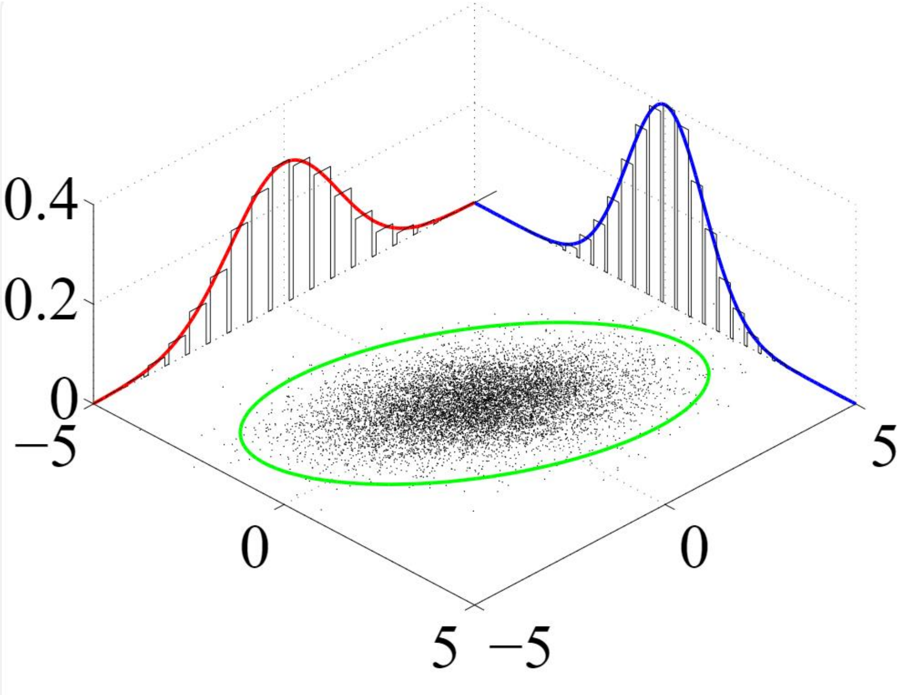
Marginal density function

When we only consider part of the variables.

X is a random vector of size n. We consider 'sub-vector' of size k (k < n).

Marginal density function: $f_S(x_1, \ldots, x_k) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f(x_1, \ldots, x_n)\, dx_{k+1} \ldots dx_n$

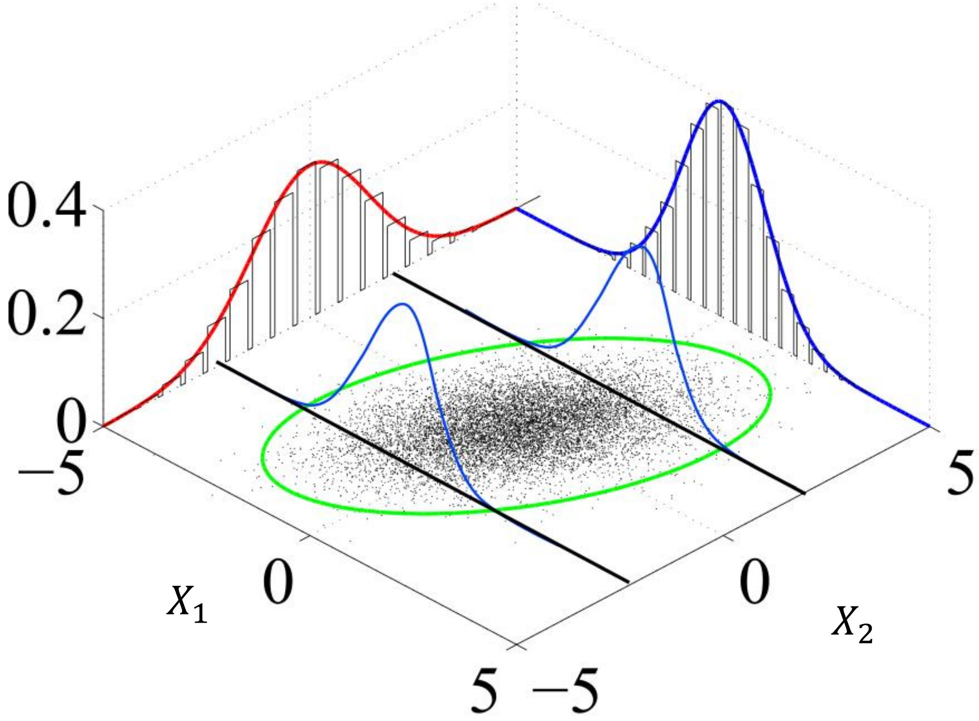
Example in 2D:



Conditional density function

When part of the variables are fixed.

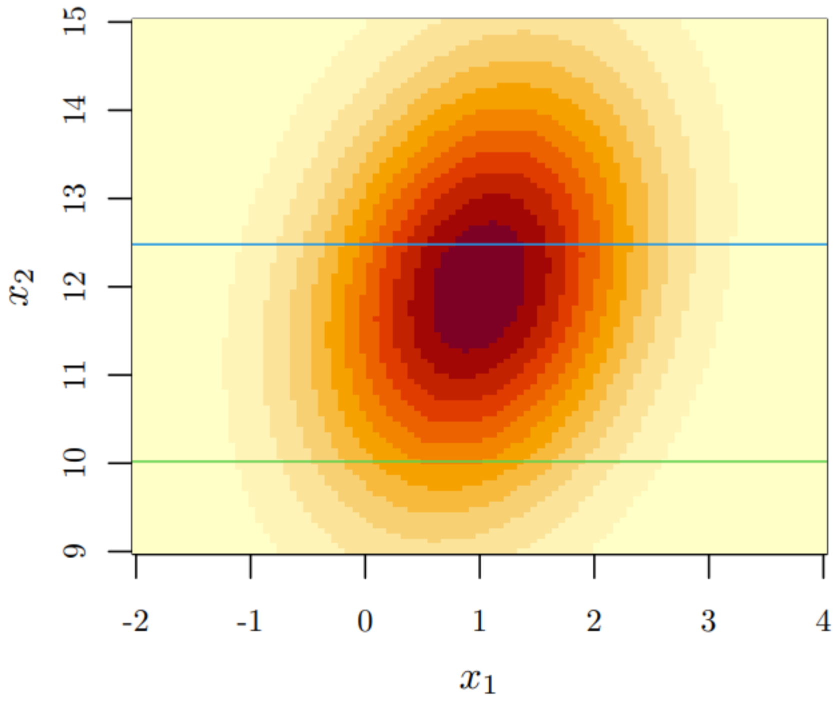
Conditional density function: $f_{X_1|X_2=x_2}(x_1) = \dfrac{f_{X_1,X_2}(x_1,x_2)}{f_{X_2}(x_2)}$

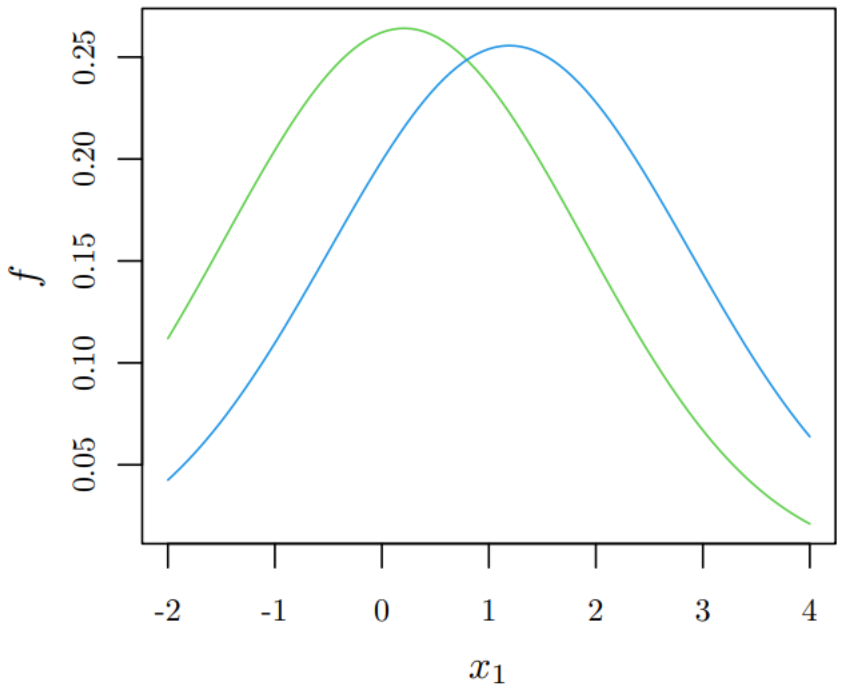


Conditional density function

When part of the variables are fixed.

Conditional density function: $$f_{X_1|X_2=x_2}(x_1) = \frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_2}(x_2)}$$



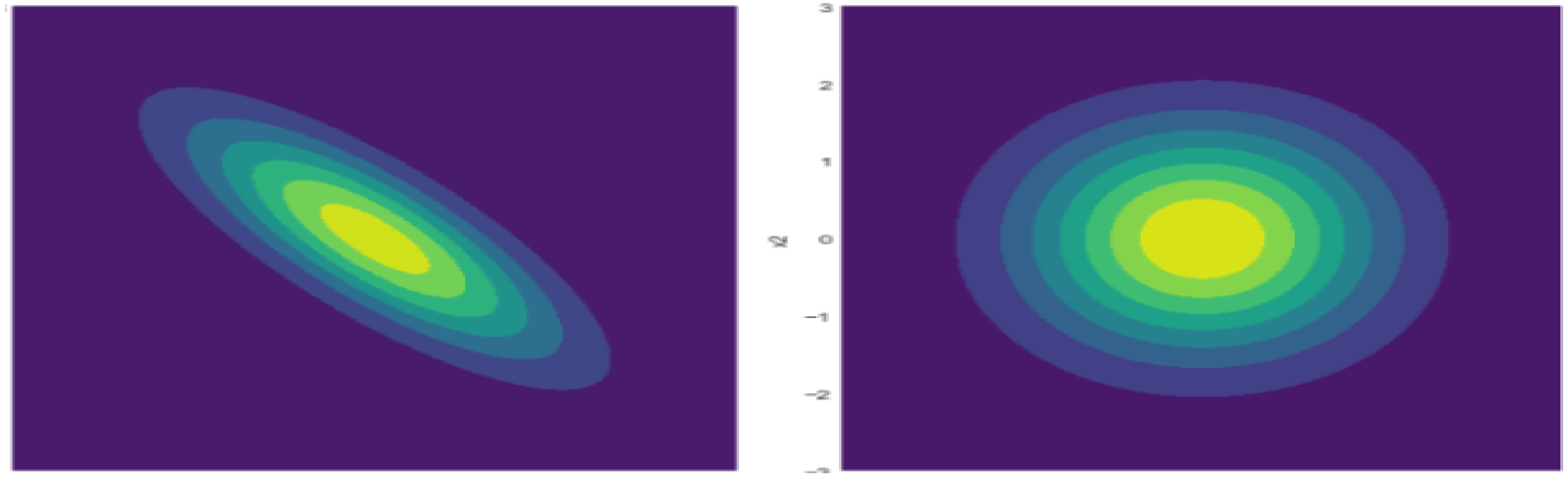


Independence

If variables X and Y are independent, then: $f_{X,Y}(x,y) = f_X(x) f_Y(y)$

If X and Y are independent, then the distribution of Y does not depend on X: $f_{Y|X=x}(y) = f_Y(y)$

Independence implies no correlation – not the other way around



Expectation

E[X] = the **expectation value** is the same as the average of unrealized random variables.

$$E[X] = \int_{-\infty}^{\infty} x\, f_X(x) dx$$

E[..] is also called the “expectation operator”

Expectation is a **linear operator**:

$$E[a + bX_1 + cX_2] = a + b\,E[X_1] + c\,E[X_2]$$

Moments and central moments

n’th moment: $$E[X^n] = \int_{-\infty}^{\infty} x^n f_X(x)\, dx$$

n’th central moment: $$E[(X - E[X])^n] = \int_{-\infty}^{\infty} (x - E[X])^n f_X(x)\, dx$$

Variance and covariance

The variance is the same as the 2nd central moment:

$$V[X] = E[(X - E[X])^2] = E[X^2] - (E[X])^2$$

Covariance is a ***mixed*** central moment:

$$\text{Cov}[X_1, X_2] = E[(X_1 - E[X_1])(X_2 - E[X_2])] = E[X_1 X_2] - E[X_1]E[X_2]$$

$$V[X] = \text{Cov}[X, X]$$

Moments of random vectors

Expectation: $E[\boldsymbol{X}] = [E[X_1], E[X_2], \ldots, E[X_n]]^T$

Variance-covariance matrix:

$$\boldsymbol{\Sigma}_X = V[\boldsymbol{X}] = E[(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})^T] =$$

$$\begin{bmatrix} V[X_1] & \text{Cov}[X_1, X_2] & \cdots & \text{Cov}[X_1, X_n] \\ \text{Cov}[X_2, X_1] & V[X_2] & \cdots & \text{Cov}[X_2, X_n] \\ \vdots & & & \vdots \\ \text{Cov}[X_n, X_1] & \text{Cov}[X_n, X_2] & \cdots & V[X_n] \end{bmatrix}$$

Correlation: $\rho_{ij} = \dfrac{\text{Cov}[X_i, X_j]}{\sqrt{V[X_i]V[X_j]}} = \dfrac{\sigma_{ij}}{\sigma_i \sigma_j}$

Calculation rule

Calculation rule worth remembering (2.28):

$$\text{Cov}[aX_1 + bX_2, cX_3 + dX_4] =$$
$$ac\,\text{Cov}[X_1, X_3] + ad\,\text{Cov}[X_1, X_4] + bc\,\text{Cov}[X_2, X_3] + bd\,\text{Cov}[X_2, X_4]$$

The rule can be used for variance as well:

$$V[a + bX_2] = b^2 V[X_2]$$
$$V[aX_1 + bX_2] = a^2 V[X_1] + b^2 V[X_2] + 2ab\text{Cov}[X_1, X_2]$$

Moment representation

With the “moment representation” we describe all the moments up to a certain order (but we do not describe the full probability distribution).

“Second order moment representation” consists of:

- Mean (expectation)
- Variance
- Covariance (if relevant)

The normal distribution is fully characterized by its second order representation.

Conditional expectations

Conditional expectation: $$\mathrm{E}[Y|X=x] = \int_{-\infty}^{\infty} y f_{Y|X=x}(y)\, dy$$

Conditional variance: $$\mathrm{Var}[\boldsymbol{Y}|\boldsymbol{X}] = \mathrm{E}[(\boldsymbol{Y} - \mathrm{E}[\boldsymbol{Y}|\boldsymbol{X}])(\boldsymbol{Y} - \mathrm{E}[\boldsymbol{Y}|\boldsymbol{X}])^T|\boldsymbol{X}]$$

Variance separation theorem ((2.51) and (2.52) in the book):

$$\mathrm{Var}[\boldsymbol{Y}] = \mathrm{E}[\mathrm{Var}[\boldsymbol{Y}|\boldsymbol{X}]] + \mathrm{Var}[\mathrm{E}[\boldsymbol{Y}|\boldsymbol{X}]]$$
$$\mathrm{C}[\boldsymbol{Y},\boldsymbol{Z}] = \mathrm{E}[\mathrm{C}[\boldsymbol{Y},\boldsymbol{Z}|\boldsymbol{X}]] + \mathrm{C}[\mathrm{E}[\boldsymbol{Y}|\boldsymbol{X}],\mathrm{E}[\boldsymbol{Z}|\boldsymbol{X}]]$$

Obs: in the last equation C[..] is the same as Cov[..]

Exercises

Exercise 2.1, 2.2, 2.3

Use (2.25) (2.28) (2.39) (2.40) (2.51) and (2.52) from the book

Obs: Ex 2.2 should say “independent” instead of “uncorrelated”