

## Course 02402 Introduction to Statistics

### Lecture 13: A course summary

DTU Compute  
Technical University of Denmark  
2800 Lyngby – Denmark

## Agenda - the 12 lectures

- 1 Chapter 1: Simple Graphics and Summary Statistics
- 2 Chapter 2: Discrete Distributions
- 3 Chapter 2: Continuous Distributions
- 4 Chapter 3: One sample confidence intervals
- 5 Chapter 3: One sample hypothesis testing
- 6 Chapter 3: Two Sample statistics
- 7 Chapter 4: Statistics by simulation
- 8 Chapter 5: Simple linear Regression Analysis
- 9 Chapter 6: Multiple linear Regression Analysis
- 10 Chapter 8: One-way Analysis of Variance
- 11 Chapter 8: Two-way Analysis of Variance
- 12 Chapter 7: Inferences for Proportions

## Overview

- 1 Chapter 1: Simple Graphics and Summary Statistics
- 2 Chapter 2: Discrete Distributions
- 3 Chapter 2: Continuous Distributions
- 4 Chapter 3: One sample confidence intervals
- 5 Chapter 3: One sample hypothesis testing
- 6 Chapter 3: Two Sample statistics
- 7 Chapter 4: Statistics by simulation
- 8 Chapter 5: Simple linear Regression Analysis
- 9 Chapter 6: Multiple linear Regression Analysis
- 10 Chapter 8: One-way Analysis of Variance
- 11 Chapter 8: Two-way Analysis of Variance
- 12 Chapter 7: Inferences for Proportions

## Chapter 1: Simple Graphics and Summary Statistics

- Look at data as it is! (descriptive statistics)
- Summary Statistics
  - Mean  $\bar{x}$
  - Standard deviation  $s$ , variance  $s^2$
  - Median, upper- and lower quartiles
- Simple graphics
  - Scatter plot (xy plot)
  - Histogram, cumulative distribution
  - Boxplots, Bar charts, Pie charts

## Overview

- 1 Chapter 1: Simple Graphics and Summary Statistics
- 2 Chapter 2: Discrete Distributions**
- 3 Chapter 2: Continuous Distributions
- 4 Chapter 3: One sample confidence intervals
- 5 Chapter 3: One sample hypothesis testing
- 6 Chapter 3: Two Sample statistics
- 7 Chapter 4: Statistics by simulation
- 8 Chapter 5: Simple linear Regression Analysis
- 9 Chapter 6: Multiple linear Regression Analysis
- 10 Chapter 8: One-way Analysis of Variance
- 11 Chapter 8: Two-way Analysis of Variance
- 12 Chapter 7: Inferences for Proportions

## Chapter 2: Discrete Distributions

- General concepts:
  - Definition of a stochastic variable
  - Density function
  - Distribution function
  - Mean and variance
- Specific distributions:
  - The binomial distribution
  - The hypergeometric distribution
  - The Poisson distribution

## Overview

- 1 Chapter 1: Simple Graphics and Summary Statistics
- 2 Chapter 2: Discrete Distributions
- 3 Chapter 2: Continuous Distributions**
- 4 Chapter 3: One sample confidence intervals
- 5 Chapter 3: One sample hypothesis testing
- 6 Chapter 3: Two Sample statistics
- 7 Chapter 4: Statistics by simulation
- 8 Chapter 5: Simple linear Regression Analysis
- 9 Chapter 6: Multiple linear Regression Analysis
- 10 Chapter 8: One-way Analysis of Variance
- 11 Chapter 8: Two-way Analysis of Variance
- 12 Chapter 7: Inferences for Proportions

## Chapter 2: Continuous Distributions

- General concepts:
  - Density function, distribution function
  - Mean, variance
  - Calculation rules for stochastic variables
- Specific distributions:
  - Normal
  - Log-Normal, Uniform, Exponential

## Overview

- 1 Chapter 1: Simple Graphics and Summary Statistics
- 2 Chapter 2: Discrete Distributions
- 3 Chapter 2: Continuous Distributions
- 4 Chapter 3: One sample confidence intervals**
- 5 Chapter 3: One sample hypothesis testing
- 6 Chapter 3: Two Sample statistics
- 7 Chapter 4: Statistics by simulation
- 8 Chapter 5: Simple linear Regression Analysis
- 9 Chapter 6: Multiple linear Regression Analysis
- 10 Chapter 8: One-way Analysis of Variance
- 11 Chapter 8: Two-way Analysis of Variance
- 12 Chapter 7: Inferences for Proportions

## Chapter 3: One sample confidence intervals

- General concepts
  - Estimation, confidence intervals
  - Population and a random sample
  - Sampling distributions ( $t$  and  $\chi^2$ )
  - Central Limit Theorem
- Specific methods, one sample:
  - Confidence intervals for the mean
  - Confidence intervals for the variance (and standard deviation)

## Overview

- 1 Chapter 1: Simple Graphics and Summary Statistics
- 2 Chapter 2: Discrete Distributions
- 3 Chapter 2: Continuous Distributions
- 4 Chapter 3: One sample confidence intervals
- 5 Chapter 3: One sample hypothesis testing**
- 6 Chapter 3: Two Sample statistics
- 7 Chapter 4: Statistics by simulation
- 8 Chapter 5: Simple linear Regression Analysis
- 9 Chapter 6: Multiple linear Regression Analysis
- 10 Chapter 8: One-way Analysis of Variance
- 11 Chapter 8: Two-way Analysis of Variance
- 12 Chapter 7: Inferences for Proportions

## Chapter 3: One sample hypothesis testing

- General concepts:
  - Hypotheses, p-value, Significance level
  - Type I and Type II error, Power
- Specific methods, One sample:
  - $t$ -test for mean difference
  - Sample size for wanted power
  - Normal qq-plot

## Overview

- 1 Chapter 1: Simple Graphics and Summary Statistics
- 2 Chapter 2: Discrete Distributions
- 3 Chapter 2: Continuous Distributions
- 4 Chapter 3: One sample confidence intervals
- 5 Chapter 3: One sample hypothesis testing
- 6 Chapter 3: Two Sample statistics**
- 7 Chapter 4: Statistics by simulation
- 8 Chapter 5: Simple linear Regression Analysis
- 9 Chapter 6: Multiple linear Regression Analysis
- 10 Chapter 8: One-way Analysis of Variance
- 11 Chapter 8: Two-way Analysis of Variance
- 12 Chapter 7: Inferences for Proportions

## Chapter 3: Two Samples

- Specific methods, two samples:
  - Test and confidence interval for the mean difference ( $t$ -test)
- Specific methods, two PAIRED samples:
  - "Take difference"  $\Rightarrow$  "One sample"
- Planning for precision and/or power
  - One-sample Confidence interval: sample size for wanted precision
  - One-sample hypothesis test: sample size for wanted power (or other combinations)
  - Two-sample hypothesis test: sample size for wanted power (or other combinations)

## Overview

- 1 Chapter 1: Simple Graphics and Summary Statistics
- 2 Chapter 2: Discrete Distributions
- 3 Chapter 2: Continuous Distributions
- 4 Chapter 3: One sample confidence intervals
- 5 Chapter 3: One sample hypothesis testing
- 6 Chapter 3: Two Sample statistics
- 7 Chapter 4: Statistics by simulation**
- 8 Chapter 5: Simple linear Regression Analysis
- 9 Chapter 6: Multiple linear Regression Analysis
- 10 Chapter 8: One-way Analysis of Variance
- 11 Chapter 8: Two-way Analysis of Variance
- 12 Chapter 7: Inferences for Proportions

## Chapter 4, Statistics by simulation

- Introduction to simulation
- Error propagation rules
- Bootstrapping
  - Parametric
  - Non-parametric
  - Confidence intervals (and hence also hypothesis testing)
- Specific situations: (4 versions of confidence intervals)
  - One-sample and Two-sample data
  - Parametric and Non-parametric

## Overview

- 1 Chapter 1: Simple Graphics and Summary Statistics
- 2 Chapter 2: Discrete Distributions
- 3 Chapter 2: Continuous Distributions
- 4 Chapter 3: One sample confidence intervals
- 5 Chapter 3: One sample hypothesis testing
- 6 Chapter 3: Two Sample statistics
- 7 Chapter 4: Statistics by simulation
- 8 Chapter 5: Simple linear Regression Analysis**
- 9 Chapter 6: Multiple linear Regression Analysis
- 10 Chapter 8: One-way Analysis of Variance
- 11 Chapter 8: Two-way Analysis of Variance
- 12 Chapter 7: Inferences for Proportions

## Chapter 5: Simple linear Regression Analysis

- Two quantitative variables,  $x$  and  $y$ .
- Calculating least squares line
- Inferences for a simple linear regression model
  - Statistical model:  $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$
  - Interval estimation and test for  $\beta_0$  and  $\beta_1$ .
  - Confidence interval for the expected line.
  - Prediction interval.
- $r$  and  $r^2$ 
  - $r$  describes the strength of a linear relation.
  - $r^2$  expresses the proportion of the  $y$  variability explained by the linear relation.

## Overview

- 1 Chapter 1: Simple Graphics and Summary Statistics
- 2 Chapter 2: Discrete Distributions
- 3 Chapter 2: Continuous Distributions
- 4 Chapter 3: One sample confidence intervals
- 5 Chapter 3: One sample hypothesis testing
- 6 Chapter 3: Two Sample statistics
- 7 Chapter 4: Statistics by simulation
- 8 Chapter 5: Simple linear Regression Analysis
- 9 Chapter 6: Multiple linear Regression Analysis**
- 10 Chapter 8: One-way Analysis of Variance
- 11 Chapter 8: Two-way Analysis of Variance
- 12 Chapter 7: Inferences for Proportions

## Chapter 6: Multiple linear Regression Analysis

- Many quantitative variables,  $x_1$ ,  $x_2$  and  $y$ .
- Calculating least squares fit
- Inferences for a the multiple linear regression model
  - Statistical model:  $y_i = \beta_0 + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \varepsilon_i$
  - Interval estimation and test for  $\beta_0$  and  $\beta_i$ .
  - Confidence interval for the expected fit.
  - Prediction interval.
- $r^2$  expresses the proportion of the  $y$  variability explained by the linear relation.

## Overview

- 1 Chapter 1: Simple Graphics and Summary Statistics
- 2 Chapter 2: Discrete Distributions
- 3 Chapter 2: Continuous Distributions
- 4 Chapter 3: One sample confidence intervals
- 5 Chapter 3: One sample hypothesis testing
- 6 Chapter 3: Two Sample statistics
- 7 Chapter 4: Statistics by simulation
- 8 Chapter 5: Simple linear Regression Analysis
- 9 Chapter 6: Multiple linear Regression Analysis
- 10 Chapter 8: One-way Analysis of Variance**
- 11 Chapter 8: Two-way Analysis of Variance
- 12 Chapter 7: Inferences for Proportions

## Chapter 8: One-way Analysis of Variance

- Specific methods,  $k$  INDEPENDENT samples
- One-way analysis of variance
  - Compares the means of the groups
  - ANOVA-table:  $SST = SS(Tr) + SSE$
  - $F$ -test.
  - Post hoc test: pairwise  $t$ -test with/without Bonferroni correction

## Overview

- 1 Chapter 1: Simple Graphics and Summary Statistics
- 2 Chapter 2: Discrete Distributions
- 3 Chapter 2: Continuous Distributions
- 4 Chapter 3: One sample confidence intervals
- 5 Chapter 3: One sample hypothesis testing
- 6 Chapter 3: Two Sample statistics
- 7 Chapter 4: Statistics by simulation
- 8 Chapter 5: Simple linear Regression Analysis
- 9 Chapter 6: Multiple linear Regression Analysis
- 10 Chapter 8: One-way Analysis of Variance
- 11 Chapter 8: Two-way Analysis of Variance**
- 12 Chapter 7: Inferences for Proportions

## Chapter 8: Two-way Analysis of Variance

- Block design - two-way analysis of variance
- ANOVA-table:  $SST = SS(Tr) + SS(BI) + SSE$ 
  - $SST$ ,  $SS(Tr)$  and  $SS(BI)$  calculated as one-way ANOVA
  - $SSE = SST - SS(Tr) - SS(BI)$
- $F$ -test.
- Post hoc test: pairwise  $t$ -test with/without Bonferroni correction

## Overview

- 1 Chapter 1: Simple Graphics and Summary Statistics
- 2 Chapter 2: Discrete Distributions
- 3 Chapter 2: Continuous Distributions
- 4 Chapter 3: One sample confidence intervals
- 5 Chapter 3: One sample hypothesis testing
- 6 Chapter 3: Two Sample statistics
- 7 Chapter 4: Statistics by simulation
- 8 Chapter 5: Simple linear Regression Analysis
- 9 Chapter 6: Multiple linear Regression Analysis
- 10 Chapter 8: One-way Analysis of Variance
- 11 Chapter 8: Two-way Analysis of Variance
- 12 Chapter 7: Inferences for Proportions

## Chapter 7: Inferences for Proportions

- Specific methods, one, two and  $k > 2$  samples
  - Binary/categorical response
- Estimation and confidence interval of proportions
  - Large sample vs. small sample methods
- Hypotheses for one proportion
- Hypotheses for two proportions
- Analysis of contingency tables ( $\chi^2$ -test) (All expected  $> 5$ )

## Agenda

- 1 Chapter 1: Simple Graphics and Summary Statistics
- 2 Chapter 2: Discrete Distributions
- 3 Chapter 2: Continuous Distributions
- 4 Chapter 3: One sample confidence intervals
- 5 Chapter 3: One sample hypothesis testing
- 6 Chapter 3: Two Sample statistics
- 7 Chapter 4: Statistics by simulation
- 8 Chapter 5: Simple linear Regression Analysis
- 9 Chapter 6: Multiple linear Regression Analysis
- 10 Chapter 8: One-way Analysis of Variance
- 11 Chapter 8: Two-way Analysis of Variance
- 12 Chapter 7: Inferences for Proportions