Module 3 - Statistical Inference Part I

1. Introduction to hypothesis testing and confidence intervals
2. One-sample Student's test and confidence interval
3. Paired data, two independent samples

Valeria Vitelli
Oslo Centre for Biostatistics and Epidemiology
Department of Biostatistics, UiO
valeria.vitelli@medisin.uio.no

MF9130E – Introductory Course in Statistics
26.04.2023

Overview for Module 3, part I

Schedule for today: Lectures in flipped classroom style (FOR*)

- 12:45 – 13:15 Introductory lecture
- 13:15 – 14:15 Self study session
- 14:15 – 14:30 Break
- 14:30 – 15:15 Group work session
- 15:15 – 15:30 Closing session (Wrap-up / Q&A)

Tomorrow morning: Lab in flipped classroom style (SEM*)

- One-sample t-Student test and confidence interval with **R**
- Two-sample t-Student test and confidence interval with **R**

Introductory Lecture for Module 3, part I

- Shortly explain the **key concepts** for today
- Make clear **what to focus on**
- Explain how to use the **learning material** in the **Self study session**
- Give **guiding questions** for the **Group work session**

Introductory Lecture for Module 3, part I

Key concepts

1. **Properties of the Sample Mean**
   - Recap from Module 1
   - Standard deviation vs. standard error
2. **Confidence Intervals (CIs)**
   - The Student t-distribution
   - Confidence interval for the mean
3. **Testing an hypothesis**
   - One-sample t-test
   - Paired data
   - Two sample t-test

Key concept 1. Properties of the Sample Mean

Recap from Module 1

How to describe the data distribution?

- **central measures** (mean, median, mode)
- **measures of variation** (range / IQR, empirical variance / standard deviation)

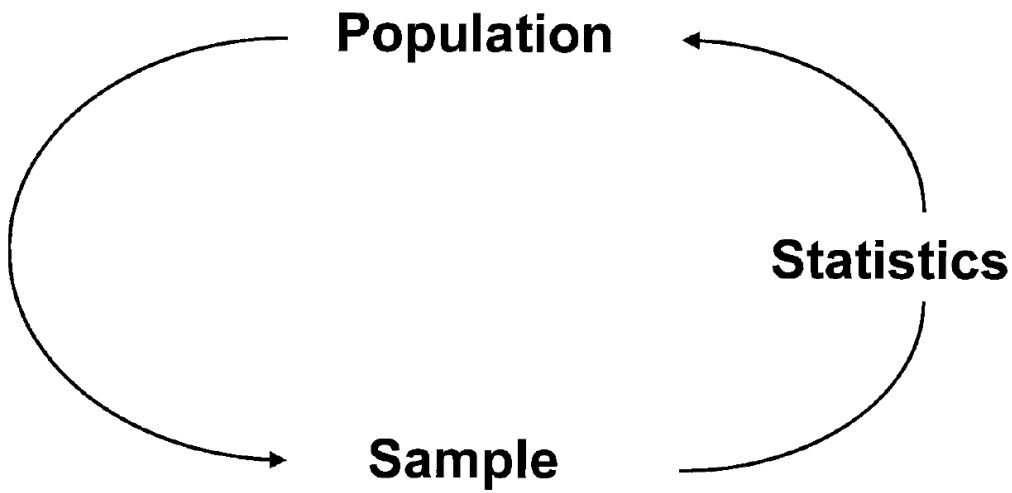
Some facts:

- which measure to use depends on the situation
- the sample mean $\bar{X}$ is an estimator for the **population** mean μ
- the empirical standard deviation s is an estimator for the population standard deviation σ
- the sample mean $\bar{X}$ is a **normally distributed random variable**, with mean μ and standard deviation $\sigma/\sqrt{n}$, the latter is also called **standard error**

Key concept 1. Properties of the Sample Mean

Recap from Module 1

- **Inferential Statistics** is about using information from a sample (data set) to **make inference** about the population it originates from
- Therefore, the **sample** is of interest for what it tells the investigator about the **population** which it represents



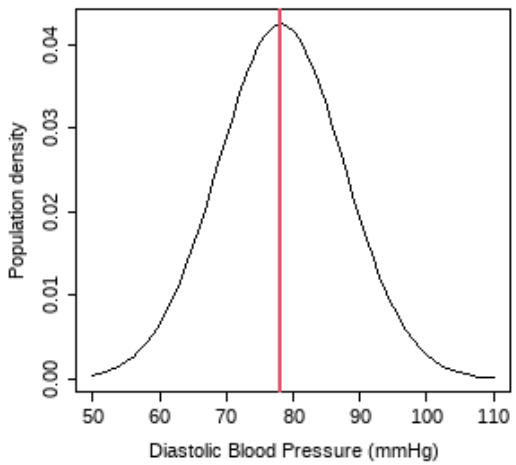
Key concept 1. Properties of the Sample Mean

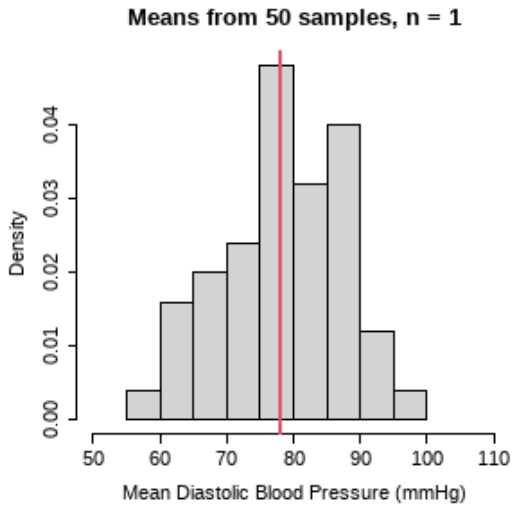
standard deviation vs. standard error

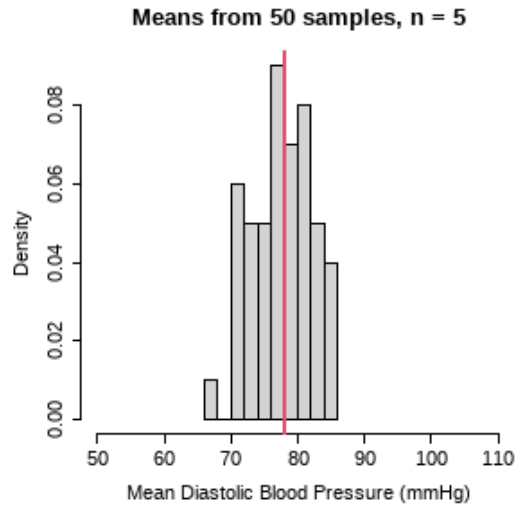
- every sample will give a different estimate of $\bar{X}$ due to **sample variation**
- the **standard error** of the sample mean reflects this variation, as it measures *how precisely the population mean μ is estimated by the sample mean $\bar{X}$*
- **by construction,** the standard error decreases when the sample size n increases (also natural / intuitive!).

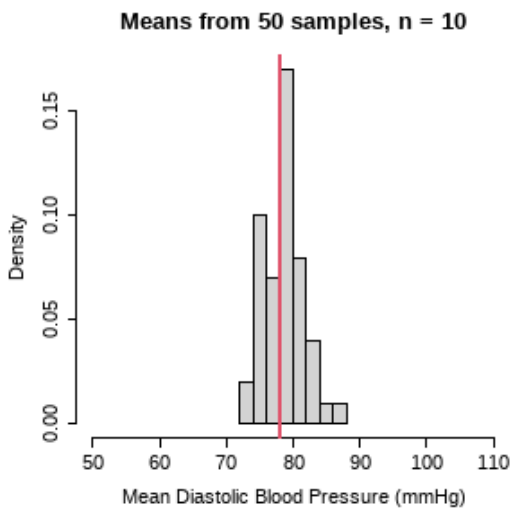
Key concept 1. standard deviation vs. standard error

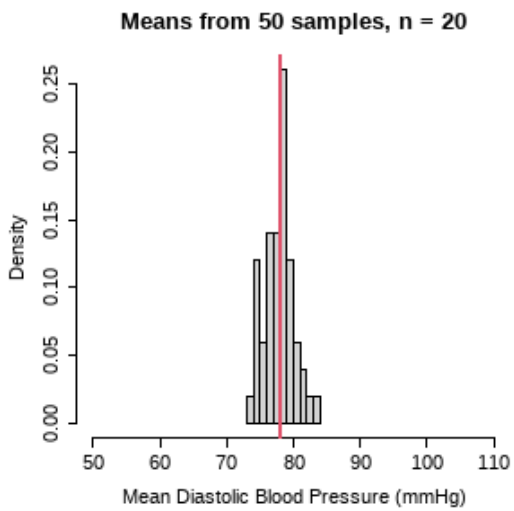
Example 4.4 at page 39-41 of K&S $\rightarrow$ **simulated in R!**

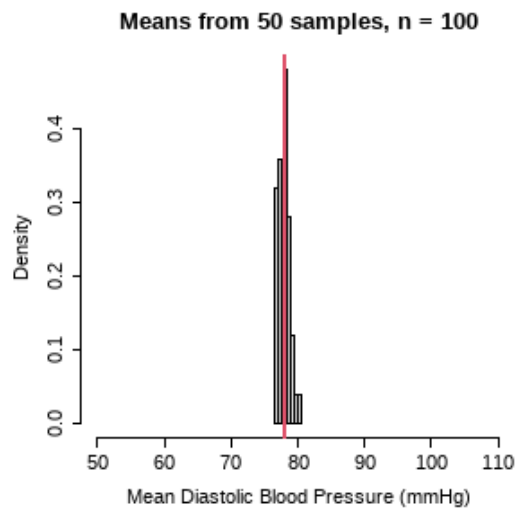








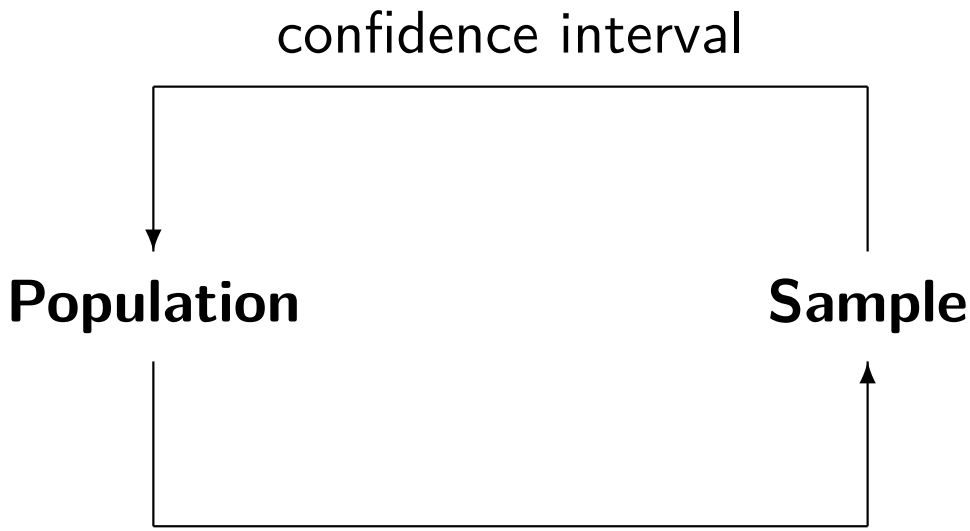




Key concept 2. Confidence Intervals (CIs)

Confidence intervals for the mean

- The sample mean $\bar{X}$ is an **estimate** of the true mean μ in the whole population
- We seek a method to **quantify how representative our estimate is**
- We are able to construct a **range of likely values**, called a **confidence interval** (CI), for the (unknown) population mean based on the sample mean and its standard error



Key concept 2. Confidence Intervals (CIs)

95%-confidence interval

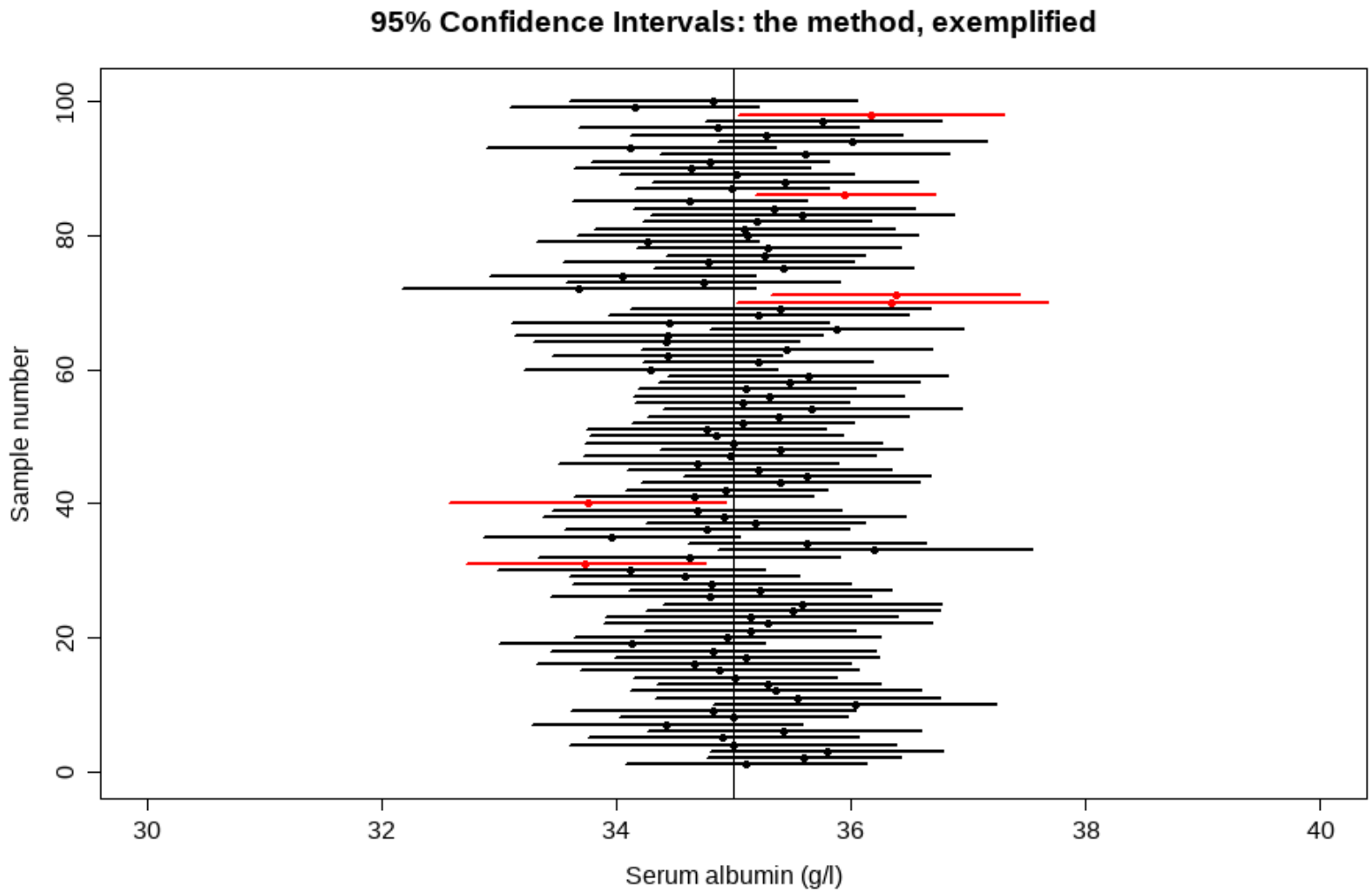
- A method that we can apply to the sample to produce an interval
- The probability that this method will produce an interval that contains the true value is 95%
- We will refer to such interval as 95% *confidence interval*

common misunderstanding

this is not the same as saying that our estimated interval contains the true value with 95% probability!

Key concept 2. Confidence Intervals (CIs)

Simulation in R: Confidence intervals for the **mean serum albumin** constructed from 100 random samples of size $n = 25$. Vertical line at the population mean μ; red CIs do not cover μ



Key concept 3. Testing an hypothesis

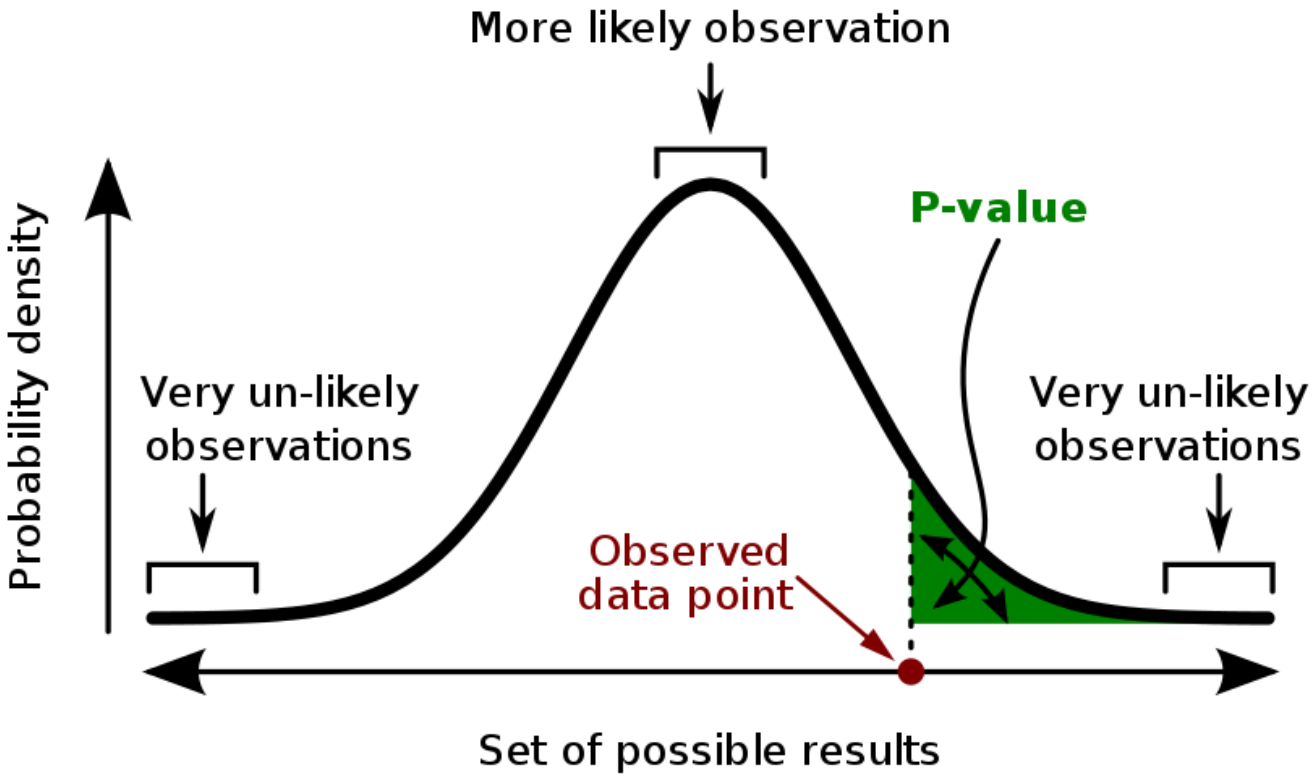
Hypothesis testing in general

- State your **null hypothesis** H_0: **aim** of the test is to check whether the data provide **sufficient evidence to reject it**
- Derive the **test statistic**, who has a certain distribution
- **Take a decision**, accept/reject, or calculate the **p-value**. (By definition, if the p-value is below a certain level you can reject H_0)

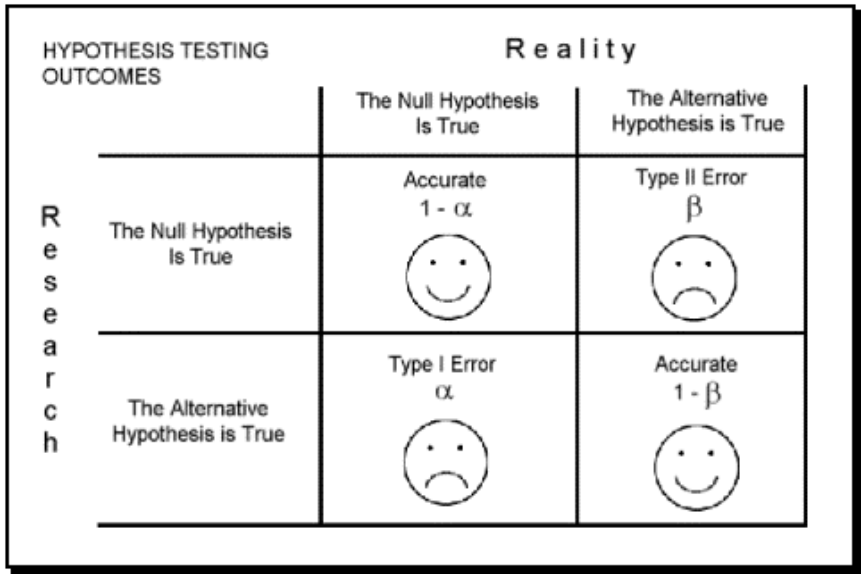
P-value: definition

The probability that the observed result, or a result more extreme, is true, given H_0 is true.

P-value


More likely observation
P-value
Probability density
Very un-likely observations
Very un-likely observations
Observed data point
Set of possible results

Type I and Type II errors



- **Type I error** (false positive):
  $P(H_0 \text{ rejected} \mid H_0 \text{ true}) = \alpha$
  Also called **level** of the test, as it defines the test itself
  (α is thus determined in advance, example value 5%)
- **Type II error** (false negative):
  $P(H_0 \text{ not rejected} \mid H_0 \text{ false}) = \beta$
  Influenced by sample size; it is equal to 1 - **power**

Test procedure

1. Formulate the test (null hypothesis & alternative hypothesis)
2. Choose an appropriate test and level α
3. Calculate the test-statistic
4. - Compare the test-statistic with the α-threshold, **OR**
   - Calculate the p-value, and compare it with α
5. Decide whether the null hypothesis is to be rejected or not
6. Formulate the conclusion

Correspondence between CI and test

If the 95%-CI for μ does not include μ_0, then the corresponding test can be rejected at the 5% level

Summary

Key terms and concepts

- Recap from Module 1 (concepts from **Descriptive Statistics**)
- **Inferential Statistics** (as opposed to descriptive statistics)
- **Population** and **sample**, properties of the **sample mean**
- standard deviation vs **standard error** of the mean
- **Confidence intervals:** general idea, concept of **coverage**
  - CI for the mean when σ is known
  - CI for the mean when σ is unknown, the Student t-distribution
  - Non-normality, small sample sizes
  - CI for the mean difference $\mu_1 - \mu_0$ of two independent samples
- **Testing an hypothesis:** general idea, concept of **p-value**
  - one-sample t-test for the population mean
  - test for paired data
  - two-independent-samples t-test for the mean difference $\mu_1 - \mu_0$

Self study session – Tasks

1. **Deepen your understanding** of each **key concept** from the previous slides by reading the corresponding longer slides:
   - `Module3-PartI-key_concept_1.pdf`
   - `Module3-PartI-key_concept_2.pdf`
   - `Module3-PartI-key_concept_3.pdf`
2. **Verify your learning outcome:**
   - **Review the Summary** (slide 16, "Key terms and concepts") in this presentation, and make sure you understand all terms
   - **IF** you feel you are still not familiar with any terms and concepts from the summary slides, use the provided **Learning Material** for this course module (next slide) to read up more
3. **Prepare for the group work session** by keeping in mind the "Guiding questions for the group work session" (slide 19) when reviewing the material

Self study session

Learning Material

- Properties of **the mean**: Aalen chapter 8.1, Kirkwood and Sterne (K&S) chapter 4
- **CIs for the mean**: Aalen chapter 8.3, K&S chapter 6 (Student t-distribution: Aalen chapter 8.2)
- One sample **t-test**: Aalen chapter 8.4
- Paired data: Aalen chapter 8.5, K&S chapter 7
- Two sample t-test: Aalen chapter 8.6, K&S chapter 7
- **General discussions** on the use of p-values and confidence intervals for interpreting results: K&S chapter 8

Group work session

Task

In your group (which should include 4-6 participants), jointly **revise the following guiding questions and provide an answer**

Guiding questions

1. What is the property of the sample mean that allows us to build Confidence Intervals and Hypothesis Tests?
2. Which is the relationship between a Confidence Interval and a Hypothesis Test, and their respective purpose?
3. The size of a p-value depends on the sample size n. How can this affect the interpretation of the p-value itself, and therefore of the analysis results?