

Inferential Statistics

Lecture 6: Estimation and confidence intervals

Jérémy Hervelin

Pforzheim University
`j.hervelin@lb.hs-pforzheim.de`

2 November 2026 15:30–17:00
Winter semester 2026/27

Adapted from the teaching materials of Prof. Dr. Thilo Klein.

Our 90-minute route

- First 15 minutes: retrieve prerequisites and introduce the question.
- Next 35 minutes: worked examples, with short checks after each concept.
- Next 25 minutes: paired activity and interpretation.
- Final 15 minutes: synthesis, questions and an applied discussion.

From a sample to an unknown parameter

- Estimate a population mean, proportion or model parameter.
- Distinguish an estimator (a random rule) from an estimate (the observed number).
- Use likelihood to compare parameter values.
- Construct and interpret a confidence interval when population SD is known.

A recurring programme example

Illustrative data: 60 of 100 sampled programme participants are employed after six months.

$$\hat{p} = 60/100 = 0.60.$$

This estimates the employment rate in the population represented by the sampling design.

It does not establish what would have happened without the programme.
Estimation of an outcome is different from estimation of a causal effect.

What makes an estimator useful?

- Bias: $E(\hat{\theta}) - \theta$. Unbiasedness concerns repeated sampling.
- Precision: the sampling variance $\text{Var}(\hat{\theta})$.
- Mean squared error: $E[(\hat{\theta} - \theta)^2] = \text{Var}(\hat{\theta}) + \text{Bias}(\hat{\theta})^2$.
- Consistency: convergence in probability to the target as sample size grows.

Sample variance: why divide by n minus one?

Estimating the sample mean uses one degree of freedom:

$$\sum_{i=1}^n (X_i - \bar{X}) = 0.$$

For independent identically distributed (iid) observations with finite variance,

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad \text{has } E(S^2) = \sigma^2.$$

The square root S estimates the population SD; it is not generally unbiased.

Likelihood: which parameter fits the observations?

Treat observed data as fixed and compare candidate parameters.

$$L(\theta; x_1, \dots, x_n) = \prod_{i=1}^n f_{\theta}(x_i)$$

for independent observations from a common PMF or density.

A maximum likelihood estimate maximises this expression. For continuous data this is a product of density heights, not the probability of an exact data vector.

Original urn example: likelihood is not a posterior

One urn has white-ball share 0.2; another has share 0.6. We draw a white ball.

The observation has likelihood 0.2 under the first urn and 0.6 under the second.

The second fits this observation better.

The probability that we chose the second urn also depends on how urns were selected. A likelihood alone does not supply that prior information.

Classroom check – Q5

Concept check

1–2 minutes

A likelihood of 0.6 establishes a 60% posterior probability for a model.

- A. True;
- B. False

Think individually, then compare your reasoning with a neighbour.

Lecture 6, classroom quiz Q5. Same question number as the printed quiz.

Bernoulli maximum likelihood

For s employed people among n independent participants:

$$L(p) = p^s(1 - p)^{n-s}.$$

For $0 < s < n$, differentiating the log-likelihood gives

$$\frac{s}{p} - \frac{n-s}{1-p} = 0 \quad \implies \quad \hat{p} = s/n.$$

If everyone is employed, the maximum is at $p = 1$; if nobody is employed, it is at $p = 0$.

Classroom check – Q1

Exam-style

1–2 minutes

60 of 100 participants are employed. Find the Bernoulli maximum likelihood estimate.

Think individually, then compare your reasoning with a neighbour.

Lecture 6, classroom quiz Q1. Same question number as the printed quiz.

Useful maximum likelihood estimates

Model	Estimate from iid observations
Bernoulli probability p	$\bar{x}$
Poisson mean λ	$\bar{x}$
Exponential rate λ	$1/\bar{x}$ (positive sample mean)
Normal mean μ	$\bar{x}$
Normal variance σ^2	$n^{-1} \sum_i (x_i - \bar{x})^2$
Uniform upper limit b on $[0, b]$	$\max_i x_i$

The normal variance MLE uses n , whereas the unbiased sample variance uses $n - 1$.

Classroom check – Q2

Concept check

1–2 minutes

Normal variance MLE divides squared deviations by...

- A. n ;
- B. $n-1$;
- C. $\sqrt{n}$

Think individually, then compare your reasoning with a neighbour.

Lecture 6, classroom quiz Q2. Same question number as the printed quiz.

Quantify the uncertainty of an estimate

For iid observations with population SD σ ,

$$SE(\bar{X}) = \sigma / \sqrt{n}.$$

When σ is unknown, its estimate is $s / \sqrt{n}$.

A confidence interval combines a point estimate with a margin of error. Its width reflects sampling uncertainty under the assumed model, not every possible source of error.

Derive the known-SD interval

If X_i are iid normal and σ is known,

$$P\left(-z_{1-\alpha/2} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq z_{1-\alpha/2}\right) = 1 - \alpha.$$

Rearrange to obtain the random interval

$$\left[\bar{X} - z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}, \quad \bar{X} + z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}\right].$$

Here z_q is the standard normal q quantile. Outside normal models, this may be a large-sample approximation.

Classroom check – Q3

Exam-style

1–2 minutes

Known $\sigma=400$ euros, $n=100$. Find the 95% margin of error using $z=1.96$.

Think individually, then compare your reasoning with a neighbour.

Lecture 6, classroom quiz Q3. Same question number as the printed quiz.

Worked illustration: monthly earnings

Synthetic example: $n = 100$, $\bar{x} = 1800$ euros, known population SD $\sigma = 400$ euros.

$$SE = 400/10 = 40 \text{ euros.}$$

$$95\% \text{ CI : } 1800 \pm 1.96(40) = [1721.60, 1878.40].$$

Interpretation concerns mean earnings in the represented population, not the earnings of 95% of individuals.

What 95 percent confidence means

Before sampling, the interval endpoints are random and the population mean is fixed.

Under repeated sampling and the model assumptions, 95% of intervals produced by this procedure cover the fixed mean.

Once this sample has been observed, the computed interval either contains that mean or does not. A frequentist confidence level is not a posterior probability for the fixed parameter.

Classroom check – Q4

Exam-style

1–2 minutes

A 95% confidence procedure means. . .

- A. The fixed parameter is random with probability 0.95 inside this observed interval;
- B. 95% of repeated intervals cover the fixed parameter under the model;
- C. 95% of people lie in the interval

Think individually, then compare your reasoning with a neighbour.

Lecture 6, classroom quiz Q4. Same question number as the printed quiz.

Original commuting-cost data

The original example contains ten costs:

2318, 2457, 2282, 2428, 2376, 2439, 2272, 2626, 2255, 2577.

Their sample mean is 2403 and sample SD is 127.5853.

We estimated the SD from these ten observations. Substituting it into a normal critical-value formula does not give an exact small-sample interval. Lecture 7 introduces the appropriate t interval.

Design activity: narrower intervals

- Compare 90%, 95% and 99% confidence while holding the data fixed. Which interval is widest?
- Predict what happens to margin of error if sample size is multiplied by four.
- Explain why a very narrow interval from a biased survey can still be misleading.
- Discuss whether a precisely estimated employment rate establishes programme effectiveness.

Independent practice

- Practise the five online questions for this lecture; 80% provides access to the next practice module.
- Course materials remain accessible regardless of your practice progress.
- Explain your reasoning before comparing answers.
- Review the method, assumptions and interpretation as well as the calculation.

Sources and teaching notes

Adapted from Prof. Dr. Thilo Klein's supplied Inferential Statistics materials.

Revised teaching edition: Jérémy Hervein, Pforzheim University, winter semester 2026/27.

New youth-employment, programme and interview-score examples are synthetic illustrations.