

Inferential Statistics

Lecture 5: Sampling distributions and the sample mean

Jérémy Hervelin

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

2 November 2026 13:45–15:15
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.

Learning objectives

- Derive the mean and variance of a sample mean under i.i.d. sampling.
- Distinguish standard deviation from standard error.
- Explain what the LLN and CLT do, and what they do not guarantee.
- Use and check a normal approximation with continuity correction.

The sample mean varies across samples

Assume $X_1, \dots, X_n$ are independent and identically distributed, with $E(X_i) = \mu$ and finite variance σ^2 .

$$\bar{X} = \frac{1}{n} \sum_i X_i.$$

Different samples produce different means. We study that variability to quantify estimation uncertainty.

These assumptions describe a model; they must be reconsidered for clustered or dependent observations.

Why the sample mean is unbiased

Linearity of expectation gives

$$E(\bar{X}) = \frac{1}{n} \sum_i E(X_i) = \frac{n\mu}{n} = \mu.$$

Independence is not needed for this equality when all means equal μ .

Unbiased means correct on average across repetitions. It does not mean that every realised sample mean equals μ .

Variance and standard error

Independence eliminates covariance terms:

$$\text{Var}(\bar{X}) = \frac{1}{n^2} \sum_i \text{Var}(X_i) = \frac{\sigma^2}{n}.$$

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

With unknown σ , estimate it by s , giving $s/\sqrt{n}$.

The standard deviation describes individual spread; the standard error describes the spread of an estimator.

Classroom check – Q1

Exam-style

1–2 minutes

Population SD is 12 and $n=36$ independent observations. Find $SE(\text{mean})$.

Think individually, then compare your reasoning with a neighbour.

Lecture 5, classroom quiz Q1. Same question number as the printed quiz.

How sample size improves precision

If $\sigma = 12$, then:

n	36	144	576
SE of the mean	2	1	0.5

Quadrupling n halves the standard error under the same i.i.d. model.

It does not halve the variability between individuals and does not repair selection bias.

Classroom check – Q2

Exam-style

1–2 minutes

Multiply n by 4 with unchanged population SD. By what factor is SE multiplied?

Think individually, then compare your reasoning with a neighbour.

Lecture 5, classroom quiz Q2. Same question number as the printed quiz.

Exact normal sampling distributions

If the observations themselves are i.i.d. $N(\mu, \sigma^2)$:

$$S_n = \sum_i X_i \sim N(n\mu, n\sigma^2), \quad \bar{X} \sim N(\mu, \sigma^2/n).$$

These statements hold exactly for every n .

Original commute example: daily mean 30 minutes, variance 9, independent days.
For three days, total mean/variance are 90/27; average mean/variance are 30/3.

Employment rates are sample means

Let $X_i = 1$ if person i is employed, and zero otherwise. Under independent Bernoulli(p) outcomes:

$$S = \sum_i X_i \sim \text{Bin}(n, p), \quad \hat{p} = S/n.$$

$$E(\hat{p}) = p, \quad \text{Var}(\hat{p}) = \frac{p(1-p)}{n}.$$

For a synthetic population with $p = 0.6$, $n = 100$, the SE is $\sqrt{0.24/100} \approx 0.0490$, or 4.90 percentage points.

Classroom check – Q4

Exam-style

1–2 minutes

For 10 independent fair tosses, find $P(0.4 < \text{proportion of heads} < 0.6)$.

Think individually, then compare your reasoning with a neighbour.

Lecture 5, classroom quiz Q4. Same question number as the printed quiz.

The original strict-inequality example

Toss a fair coin ten times. Let $\hat{p}$ be the fraction of heads.

$$P(0.4 < \hat{p} < 0.6) = P(4 < S < 6) = P(S = 5).$$

$$P(S = 5) = \binom{10}{5} (0.5)^{10} = 0.24609375.$$

The strict inequalities exclude four and six heads. Writing $\leq$ would answer a different question.

Law of large numbers

Under suitable conditions, for every $\varepsilon > 0$:

$$P(|\bar{X} - \mu| > \varepsilon) \longrightarrow 0 \quad (n \rightarrow \infty).$$

For i.i.d. observations with finite variance, Chebyshev gives

$$P(|\bar{X} - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}.$$

This is a limiting concentration result. It does not promise that each added observation moves the realised mean closer to μ .

Central limit theorem

For i.i.d. observations with finite μ and $0 < \sigma^2 < \infty$:

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0, 1).$$

For sufficiently large n , the sample mean can be approximated by $N(\mu, \sigma^2/n)$.

The theorem concerns the standardised mean, not a claim that individual observations become normally distributed.

Classroom check – Q3

Concept check

1–2 minutes

The CLT says the raw observations become normally distributed as n grows.

- A. True;
- B. False

Think individually, then compare your reasoning with a neighbour.

Lecture 5, classroom quiz Q3. Same question number as the printed quiz.

There is no universal sample-size threshold

How well the approximation works depends on skewness, tail behaviour and the event being approximated.

Rules such as “30 observations is enough” are not guarantees. For a rare binary outcome, even a seemingly large sample may contain very few successes.

For dependent or non-identically distributed variables, additional CLT conditions are required. Independence and finite individual variances alone are not a sufficient general theorem.

The original two-dice distribution

For the sum S of two fair independent dice:

s	2	3	4	5	6	7	8	9	10	11	12
$36P(S = s)$	1	2	3	4	5	6	5	4	3	2	1

$$E(S) = 7, \quad \text{Var}(S) = 70/12.$$

The exact probability $P(S \leq 4) = 6/36 = 1/6$. With only two dice, a normal approximation is fairly crude.

Continuity correction

A normal variable has no point masses. An integer count has mass at each integer. Represent the mass at integer k by an interval from $k - 0.5$ to $k + 0.5$.

$$P(S \leq k) \approx \Phi\left(\frac{k + 0.5 - \mu_S}{\sigma_S}\right).$$

For the two-dice example, the approximation improves from about 0.107 to 0.150, compared with the exact 0.1667. It remains an approximation.

Classroom check – Q5

Exam-style

1–2 minutes

For a normal approximation to an integer count, $P(X \leq 40)$ uses boundary...

- A. 39.5;
- B. 40;
- C. 40.5

Think individually, then compare your reasoning with a neighbour.

Lecture 5, classroom quiz Q5. Same question number as the printed quiz.

Original 100-toss example

For $S \sim \text{Bin}(100, 0.5)$, mean 50 and SD 5:

$$\begin{aligned}P(S \leq 40) \text{ exactly} &= 0.02844397, \\ \text{normal, no correction} &= \Phi(-2) = 0.02275013, \\ \text{normal, corrected} &= \Phi(-1.9) = 0.02871656.\end{aligned}$$

Activity: explain the change from 40 to 40.5 before calculating.

Small samples: compare with the exact answer

For $S \sim \text{Bin}(4, 0.5)$:

$$P(S \leq 1) = \frac{1 + 4}{16} = 0.3125.$$

The uncorrected normal approximation is 0.1587; the corrected one is 0.3085.

A good result for one event is not proof that an approximation works for every event. Prefer the exact binomial calculation when readily available.

Applied activity: what does a larger survey buy?

A programme survey estimates an employment rate. Compare samples of 100 and 400 independent respondents from the same population.

1. Predict how the standard error changes.
2. Explain why the observed employment rates may differ.
3. Explain why the larger survey can remain biased if employed people respond more often.

Next lecture: turn sampling uncertainty into confidence intervals.

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.