

Inferential Statistics

Lecture 3: Continuous distribution models

Jérémy Hervelin

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

12 October 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

- Calculate probabilities and quantiles for uniform, exponential and normal models.
- Distinguish a probability, a density and a quantile.
- Recognise chi-square, t and F distributions and their role in inference.

Uniform waiting times

The original bus example assumes buses depart every 30 minutes and arrival is uniform within that interval.

$$T \sim U(0, 30), \quad f_T(t) = 1/30 \quad (0 \leq t \leq 30).$$

Equal-length intervals have equal probabilities.

Not knowing the timetable alone does not establish uniform arrivals: this is a modelling assumption.

Uniform density and CDF

For $a < b$ and $X \sim U(a, b)$:

$$f_X(x) = \begin{cases} 1/(b-a) & a \leq x \leq b, \\ 0 & \text{otherwise,} \end{cases}$$

$$F_X(x) = \begin{cases} 0 & x < a, \\ (x-a)/(b-a) & a \leq x \leq b, \\ 1 & x > b. \end{cases}$$

The CDF accumulates the area under the constant density.

Classroom check – Q1

Exam-style

1–2 minutes

A uniform bus wait on $[0,30]$ minutes: find $P(\text{wait} \leq 20)$.

Think individually, then compare your reasoning with a neighbour.

Lecture 3, classroom quiz Q1. Same question number as the printed quiz.

Worked bus calculations

$$P(T > 10) = 1 - 10/30 = 2/3,$$

$$E(T) = (0 + 30)/2 = 15 \text{ minutes},$$

$$t_{0.9} = 0 + 0.9(30) = 27 \text{ minutes}.$$

A 90th percentile is a threshold not exceeded in 90% of outcomes.

For $U(a, b)$, $x_q = a + q(b - a)$ and $\text{Var}(X) = (b - a)^2/12$.

Counts and waiting times

A homogeneous Poisson process has rate λ events per unit of time.

$$N(t) \sim \text{Poisson}(\lambda t).$$

Waiting more than t for the first event means no event occurred by t :

$$P(T > t) = P(N(t) = 0) = e^{-\lambda t}.$$

The count $N(t)$ is discrete; the waiting time T is continuous.

The exponential model

For $T \sim \text{Exp}(\lambda)$, $\lambda > 0$:

$$f_T(t) = \lambda e^{-\lambda t}, \quad F_T(t) = 1 - e^{-\lambda t} \quad (t \geq 0).$$

Both functions are zero for $t < 0$.

$$E(T) = 1/\lambda, \quad \text{Var}(T) = 1/\lambda^2, \quad t_q = -\log(1 - q)/\lambda.$$

λ is a rate, not the expected duration. State its time unit.

Classroom check – Q2

Exam-style

1–2 minutes

A job-offer waiting time is exponential with rate 0.25 per month. What is its mean in months?

Think individually, then compare your reasoning with a neighbour.

Lecture 3, classroom quiz Q2. Same question number as the printed quiz.

Worked example: the original light bulb

Assume an exponential lifetime with mean 5,000 hours. Then $\lambda = 1/5000$ per hour.

$$P(T < 2500) = 1 - e^{-0.5} \approx 0.3935.$$

$$t_{0.9} = -5000 \log(0.1) \approx 11,512.93 \text{ hours.}$$

A long right tail makes the mean and the 90th percentile very different. The exponential assumption is illustrative, not a claim about all light bulbs.

Job-search durations and model assumptions

In a synthetic model, job offers arrive at rate 0.25 per month.

$$E(T) = 4 \text{ months}, \quad P(T > 4) = e^{-1} \approx 0.3679.$$

An exponential waiting time is memoryless:

$$P(T > s + t \mid T > s) = P(T > t).$$

Real job seekers may learn, change search intensity or face changing labour demand. These mechanisms can invalidate a constant-rate model.

A normal model describes a symmetric distribution

Write $X \sim N(\mu, \sigma^2)$ with $\sigma > 0$.

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x - \mu)^2}{2\sigma^2}\right].$$

μ determines location; σ determines spread. The support is all real numbers.

A histogram can suggest a normal approximation; it cannot prove normality. Bounded or strongly skewed outcomes require care.

Classroom check – Q4

Concept check

1–2 minutes

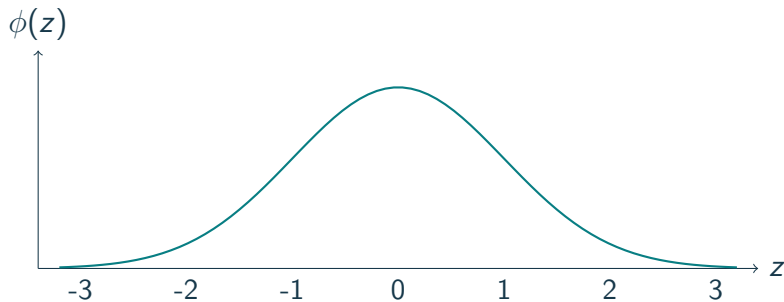
For a continuous normal variable, $P(X=a)$ equals. . .

- A. The density at a ;
- B. Zero;
- C. One

Think individually, then compare your reasoning with a neighbour.

Lecture 3, classroom quiz Q4. Same question number as the printed quiz.

The normal curve



This is $N(0, 1)$. Density height is not a point probability. Areas describe probabilities.

Standardisation

Measure the distance from the mean in standard deviations:

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1), \quad F_X(x) = \Phi\left(\frac{x - \mu}{\sigma}\right).$$

$\Phi(z) = P(Z \leq z)$ is a cumulative probability; $\phi(z)$ is a density.

Use the standard deviation σ , not the variance σ^2 , in the denominator.

Classroom check – Q3

Exam-style

1–2 minutes

Commute duration is normal with mean 40 and SD 2 minutes. What is the z score for 36 minutes?

Think individually, then compare your reasoning with a neighbour.

Lecture 3, classroom quiz Q3. Same question number as the printed quiz.

Original commute example

Assume $X \sim N(40, 4)$, in minutes. Here $\sigma = 2$.

$$P(X \leq 36) = \Phi(-2) \approx 0.02275.$$

$$P(37 \leq X \leq 43) = \Phi(1.5) - \Phi(-1.5) \approx 0.86639.$$

For a continuous model, including an endpoint makes no difference.

Activity: sketch the two areas and explain why they differ.

Normal quantiles

If $z_q = \Phi^{-1}(q)$, then $x_q = \mu + \sigma z_q$.

For the commute example:

$$x_{0.2} = 40 + 2(-0.841621) \approx 38.3168.$$

For the original exam-style model $X \sim N(15, 25)$:

$$P(X > 20) \approx 0.15866, \quad x_{0.9} \approx 21.4078.$$

The latter is a mathematical model of first-goal time. Its negative support means it is only an approximation to an actual non-negative duration.

Distributions used for test statistics

A statistic calculated from random observations is itself random.

Under particular assumptions its sampling distribution may be chi-square, t or F . We can observe a statistic's realised value; the reference distribution describes repeated samples.

Degrees of freedom are parameters of these reference distributions. They must match the statistic and the data used to estimate it.

Chi-square distribution

For independent standard normal variables $Z_1, \dots, Z_k$:

$$V = \sum_{i=1}^k Z_i^2 \sim \chi_k^2.$$

It takes non-negative values and is right-skewed, especially at small k .

Later: squared discrepancies between observed and expected categorical counts produce an approximately chi-square statistic.

Student t distribution

If $Z \sim N(0, 1)$ and $V \sim \chi_k^2$ are independent:

$$T = \frac{Z}{\sqrt{V/k}} \sim t_k.$$

The t distribution is symmetric with heavier tails than the standard normal.

$$\text{Var}(T) = \frac{k}{k-2} \quad \text{for } k > 2.$$

For $k \leq 2$ it has no finite variance. As k grows, it approaches $N(0, 1)$.

Classroom check – Q5

Concept check

1–2 minutes

Which distribution describes $(\text{mean} - \mu) / (S / \sqrt{n})$ for iid normal observations?

- A. Standard normal for every n ;
- B. t with $n-1$ degrees of freedom;
- C. Chi-square with n degrees of freedom

Think individually, then compare your reasoning with a neighbour.

Lecture 3, classroom quiz Q5. Same question number as the printed quiz.

F distribution

For independent $V \sim \chi_m^2$ and $W \sim \chi_n^2$:

$$F = \frac{V/m}{W/n} \sim F_{m,n}.$$

It is positive and generally asymmetric. The numerator and denominator degrees of freedom have different roles.

Later: comparing variances of independent normal populations.

Binomial recap and practice

For $S \sim \text{Bin}(14, 0.5)$:

$$P(S \leq 6) \approx 0.39526, \quad P(S \geq 2) = 1 - P(S \leq 1) \approx 0.99908.$$

Consider 14 independent tosses of a fair coin.

Work in pairs: explain why a binomial point probability may be positive while a normal point probability is zero.

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.