STAT 131 – Lecture Review Week 10

Normal probabilities, distribution-free bounds, the law of large numbers, and the CLT

Antonio Aguirre

Spring 2026

Based on Lectures 9.1 and 9.2, with final-style examples.

How to use this review

What this hour is for

The lecture notes give the formal statements. This review focuses on **which approximation or bound is appropriate** and how to translate a word problem into a probability statement.

Four questions before computing

1. Do we know the distribution, or only the mean and variance?
2. Are we working with a single observation, a sum, or a sample mean?
3. Is the probability one-sided, two-sided, or an interval?
4. Are we using an exact normal calculation, Chebyshev, Markov, or the CLT?

Main habit

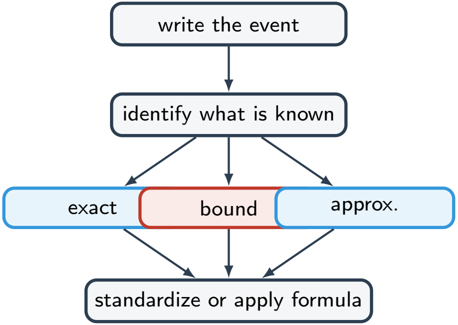
Write the event first. Then decide whether it is an exact calculation, a bound, or an approximation.

Decision map: exact, bound, or approximation?

Classify the information first

What you know	Use
Exact normal model	standardize and use Φ
Only $X \geq 0$ and $\mathbb{E}(X)$	Markov bound
Mean and variance, no distribution	Chebyshev bound
Large i.i.d. sum or mean	CLT approximation
Long-run behavior of averages	LLN statement

Key distinction: bounds are guarantees; CLT calculations are approximations.



Use this map before choosing a formula.

The standard normal is the common scale

Definition

A standard normal random variable satisfies

$$Z \sim N(0,1),$$

with density

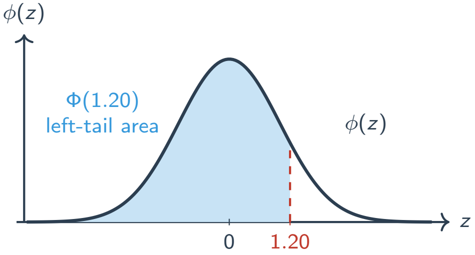
$$\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right),$$

and c.d.f.

$$\Phi(z) = \Pr(Z \le z).$$

Table interpretation

A standard normal table gives $\Phi(z)$: the left-tail area up to z.



Example lookup

$$\Phi(1.20) \approx 0.8849.$$

Normal-table workflow: left, right, and middle areas

Use the c.d.f.

For $Z \sim N(0,1)$,

$$\Pr(a < Z < b) = \Phi(b) - \Phi(a).$$

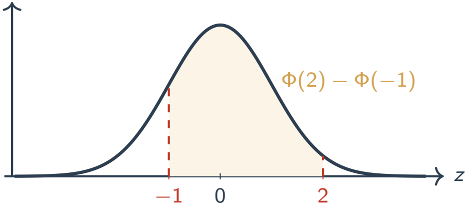
By symmetry, for $z > 0$,

$$\Phi(-z) = 1 - \Phi(z).$$

Common conversions

$$\Pr(Z > b) = 1 - \Phi(b)$$

$$\Pr(|Z| < c) = 2\Phi(c) - 1.$$



Reading tip

Tables usually give left-tail probabilities.
Convert every event into left-tail pieces.

General normal: standardize first

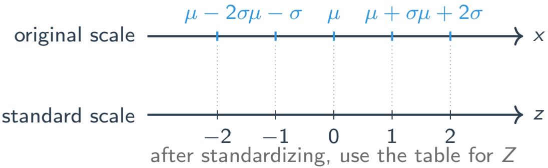
If $X \sim N(\mu, \sigma^2)$

Then

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1).$$

Therefore,

$$\Pr(a < X < b) = \Pr\left(\frac{a - \mu}{\sigma} < Z < \frac{b - \mu}{\sigma}\right).$$



Do not divide by variance

Standardize by the standard deviation σ, not by σ^2.

Quantiles

If F is the c.d.f. of $N(\mu, \sigma^2)$, then

$$F^{-1}(p) = \mu + \sigma \Phi^{-1}(p).$$

Final-style normal probability

Suppose

$$X \sim N(5, 4), \qquad \text{so } \mu = 5,\ \sigma = 2.$$

Find $\Pr(1 < X < 8)$.

Standardize both endpoints

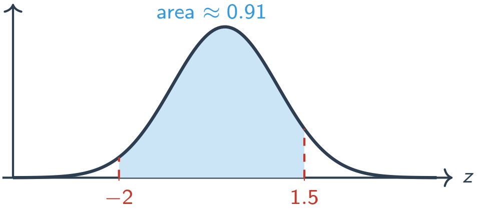
$$\Pr(1 < X < 8) = \Pr\left(\frac{1-5}{2} < Z < \frac{8-5}{2}\right)$$

$$= \Pr(-2 < Z < 1.5).$$

Use the table

$$\Pr(-2 < Z < 1.5) = \Phi(1.5) - \Phi(-2).$$

$$= 0.9332 - [1 - 0.9772] \approx 0.9104.$$



Exam habit

The first line should show the standardization, not just the final number.

Linear combinations of independent normals

Core result

If $X_1, \dots, X_n$ are independent and

$$X_i \sim N(\mu_i, \sigma_i^2),$$

then every linear combination is normal:

$$\sum_{i=1}^{n} a_i X_i \sim N\left(\sum_{i=1}^{n} a_i \mu_i,\ \sum_{i=1}^{n} a_i^2 \sigma_i^2\right).$$

Mean

Coefficients keep their sign:

$$\mathbb{E}(a_1X_1 + a_2X_2) = a_1\mu_1 + a_2\mu_2.$$

Variance

Coefficients get squared:

$$\mathrm{Var}(a_1X_1 + a_2X_2) = a_1^2\sigma_1^2 + a_2^2\sigma_2^2.$$

Independence is what allows the variance terms to add without covariance terms.

Final-style example: two bottle fills

A machine fills two independently selected bottles:

$$X_1 \sim N(16, 0.09), \qquad X_2 \sim N(16, 0.09), \qquad X_1 \perp\!\!\!\perp X_2.$$

Let

$$Y = X_1 - X_2, \qquad W = X_1 + X_2.$$

Difference

$$\mathbb{E}(Y) = 16 - 16 = 0,$$

$$\mathrm{Var}(Y) = 0.09 + 0.09 = 0.18.$$

Thus

$$Y \sim N(0, 0.18).$$

Sum

$$\mathbb{E}(W) = 16 + 16 = 32,$$

$$\mathrm{Var}(W) = 0.09 + 0.09 = 0.18.$$

Thus

$$W \sim N(32, 0.18).$$

Important

The minus sign affects the mean, but the variance contribution is still positive because $(-1)^2 = 1$.

Final-style example: using the derived distribution

Continue with

$$W = X_1 + X_2 \sim N(32, 0.18).$$

Find $\Pr(W > 32.6)$.

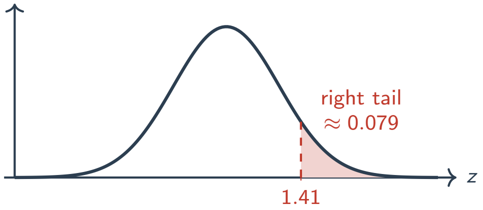
Standardize

$$\Pr(W > 32.6) = \Pr\left(Z > \frac{32.6 - 32}{\sqrt{0.18}}\right).$$

Since

$$\frac{0.6}{\sqrt{0.18}} \approx 1.41,$$

$$\Pr(W > 32.6) \approx 1 - \Phi(1.41).$$



Result

$$\Pr(W > 32.6) \approx 0.079.$$

Sample mean of normal variables

If

$$X_1, \ldots, X_n \text{ are independent}, \qquad X_i \sim N(\mu, \sigma^2),$$

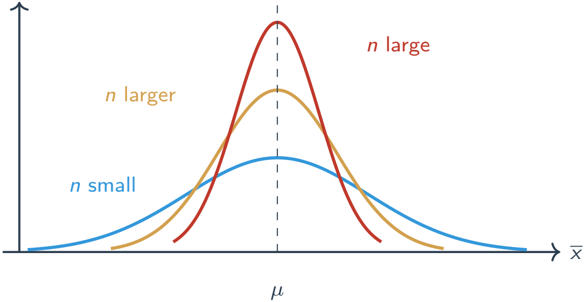
then

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

What changes with n?

$$\mathbb{E}(\overline{X}_n) = \mu, \qquad \mathrm{SD}(\overline{X}_n) = \frac{\sigma}{\sqrt{n}}.$$

The center stays fixed; the spread shrinks.



Exam habit: For a sample mean, use $\mathrm{Var}(\overline{X}_n) = \sigma^2/n$.

Final-style sample mean probability

With the bottle-fill model,

$$\overline{X}_2 = \frac{X_1 + X_2}{2}.$$

Since $X_i \sim N(16, 0.09)$,

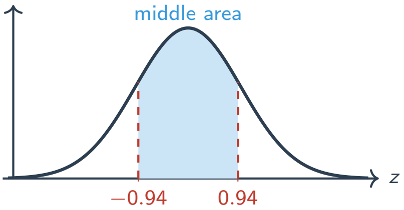
$$\overline{X}_2 \sim N\left(16, \frac{0.09}{2}\right) = N(16, 0.045).$$

Compute $\Pr(15.8 < \overline{X}_2 < 16.2)$

$$\mathrm{SD}(\overline{X}_2) = \sqrt{0.045} \approx 0.212.$$

$$\Pr(15.8 < \overline{X}_2 < 16.2) = \Pr\left(-0.94 < Z < 0.94\right).$$

$$\approx 2\Phi(0.94) - 1 \approx 0.653.$$



When we only have summary statistics

Real setting: transit wait times

Suppose a campus transit dashboard reports only summary information about the wait time X, in minutes:

$$X \geq 0, \ \mathbb{E}(X) = 12, \ \mathrm{SD}(X) = 6.$$

We do not know the full distribution of X, so we cannot compute exact probabilities. We can still give **guaranteed bounds**.

Markov: only uses $X \geq 0$ and the mean

What fraction of waits could be at least 30 minutes?

$$\Pr(X \geq 30) \leq \frac{\mathbb{E}(X)}{30} = \frac{12}{30} = 0.40.$$

This says: with only the mean, at most 40% of waits can be 30+ minutes.

Chebyshev: uses mean and variance

Since 30 is 18 minutes above the mean,

$$\{X \geq 30\} \subseteq \{|X - 12| \geq 18\}.$$

Thus

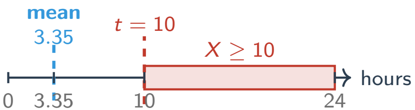
$$\Pr(X \geq 30) \leq \Pr(|X - 12| \geq 18) \leq \frac{36}{18^2} \approx 0.111$$

Markov inequality: one-sided upper tail

Theorem

If $X \geq 0$, then for $t > 0$,

$$\Pr(X \geq t) \leq \frac{\mathbb{E}(X)}{t}.$$



The picture only marks the event;
no distribution shape is assumed.

Screen-time example

Let $X =$ daily phone use, in hours. If

$$\mathbb{E}(X) = 3.35,$$

then

$$\Pr(X \geq 10) \leq \frac{3.35}{10} = 0.335.$$

Interpretation

With only $X \geq 0$ and the mean, at most 33.5% of observations can be in the (10+)-hour tail.

Source: IPA TouchPoints 2025 survey; adults in Great Britain averaged 3h21 daily mobile-phone use, reported by The Guardian.

Chebyshev inequality: two-sided distance from the mean

Theorem

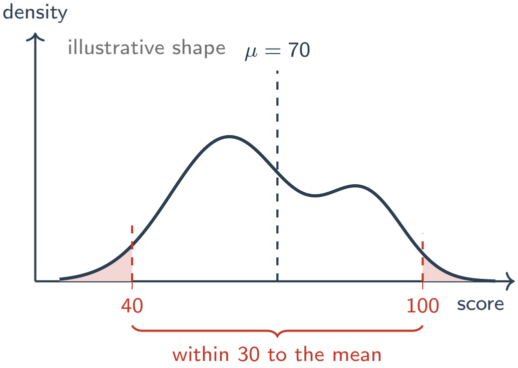
If $\mathrm{Var}(X)$ exists, then for $t > 0$,

$$\Pr(|X - \mathbb{E}(X)| \geq t) \leq \frac{\mathrm{Var}(X)}{t^2}.$$

Score example

If $\mathbb{E}(X) = 70$ and $\mathrm{SD}(X) = 10$, then

$$\Pr(|X - 70| \geq 30) \leq \frac{100}{30^2} = \frac{1}{9}.$$



Use when

The event is about distance from the mean: $|X - \mu| \geq t$ or $|X - \mu| < t$.

Final-style Chebyshev bound: set up the sample mean

Rod lengths satisfy

$$\mathbb{E}(X) = 50, \qquad \mathrm{Var}(X) = 64.$$

A quality-control engineer samples n rods and uses

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Mean and variance of the average

$$\mathbb{E}(\overline{X}_n) = 50, \qquad \mathrm{Var}(\overline{X}_n) = \frac{64}{n}.$$

What Chebyshev can do

It gives a bound without assuming the rod lengths are normal.

Bound the event

$$\Pr(|\overline{X}_n - 50| \geq 2) \leq \frac{\mathrm{Var}(\overline{X}_n)}{2^2} = \frac{64/n}{4} = \frac{16}{n}.$$

Therefore,

$$\Pr(|\overline{X}_n - 50| < 2) \geq 1 - \frac{16}{n}.$$

CLT gives a sharper approximation when justified

Compare the previous bound with a CLT calculation. Suppose

$$\overline{X}_n \approx N\left(50, \frac{64}{n}\right).$$

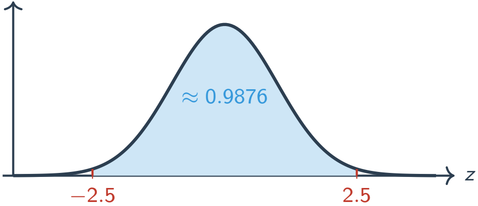
For $n = 100$,

$$\mathrm{SD}(\overline{X}_{100}) = \frac{8}{\sqrt{100}} = 0.8.$$

Approximate probability

$$\Pr(|\overline{X}_{100} - 50| < 2) \approx \Pr\left(|Z| < \frac{2}{0.8}\right)$$

$$= \Pr(|Z| < 2.5) = 2\Phi(2.5) - 1 \approx 0.9876.$$



Comparison

Chebyshev is distribution-free but conservative. CLT calculations are sharper when the approximation is justified.

Final-style CLT sample size

For the rods example, use

$$\overline{X}_n \approx N\left(50, \frac{64}{n}\right), \qquad \mathrm{SD}(\overline{X}_n) = \frac{8}{\sqrt{n}}.$$

Find the smallest n such that

$$\Pr(|\overline{X}_n - 50| < 2) \geq 0.95.$$

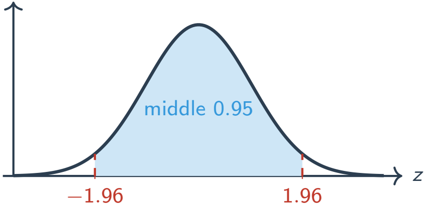
Convert to a standard normal target

$$\Pr\left(|Z| < \frac{2}{8/\sqrt{n}}\right) = \Pr\left(|Z| < \frac{\sqrt{n}}{4}\right).$$

For middle area 0.95, use $z_{0.975} \approx 1.96$:

$$\frac{\sqrt{n}}{4} \geq 1.96 \quad \Longrightarrow \quad n \geq (7.84)^2 = 61.47.$$

Thus $\boxed{n = 62}$.



Same setting, different tool

Chebyshev needs at least 320; CLT gives 62 under the normal approximation.

The Central Limit Theorem: what it says

CLT

Let $X_1, \ldots, X_n$ be i.i.d. with

$$\mathbb{E}(X_i) = \mu, \qquad \mathrm{Var}(X_i) = \sigma^2 < \infty.$$

Then, for large n,

$$\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \dot{\sim} N(0, 1).$$

Equivalently,

$$\overline{X}_n \dot{\sim} N\left(\mu, \frac{\sigma^2}{n}\right), \qquad \sum_{i=1}^{n} X_i \dot{\sim} N(n\mu, n\sigma^2).$$

Intuition

Even if the original observations are not normal, averages and large sums often behave approximately normal.

CLT visual: averages become more stable

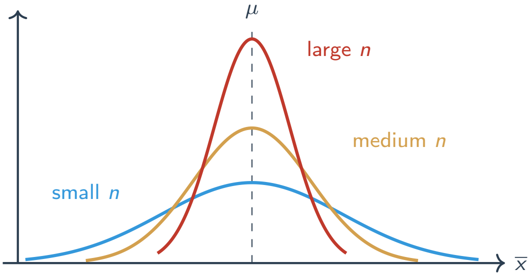
Two changes happen at once

As n grows,

- the distribution of $\overline{X}_n$ becomes more bell-shaped;
- its spread shrinks like $1/\sqrt{n}$.

Mathematical summary

$$\mathbb{E}(\overline{X}_n) = \mu, \qquad \mathrm{SD}(\overline{X}_n) = \frac{\sigma}{\sqrt{n}}.$$



Not magic

The CLT needs independent, identically distributed observations and finite variance.

CLT example: many coin tosses

A fair coin is tossed 900 times. Approximate the probability of obtaining more than 495 heads.

Exact model

Let

$$Y = \text{number of heads in 900 tosses.}$$

Then

$$Y \sim \mathrm{Binomial}(900, 0.5).$$

The exact sum is possible but long:

$$\Pr(Y > 495) = \sum_{y=496}^{900} \binom{900}{y} 0.5^{900}.$$

CLT approximation

A Bernoulli(0.5) has

$$\mu = 0.5, \qquad \sigma^2 = 0.25.$$

So

$$Y \dot{\sim} N(900 \cdot 0.5, 900 \cdot 0.25) = N(450, 225).$$

CLT example: finish the approximation

Using

$$Y \dot{\sim} N(450, 225), \qquad \mathrm{SD}(Y) = 15,$$

we approximate

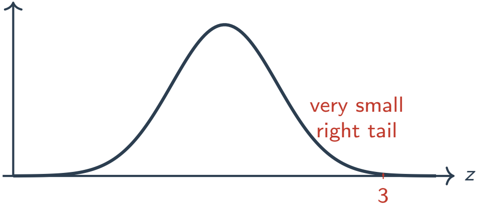
$$\Pr(Y > 495) = 1 - \Pr(Y \leq 495).$$

Standardize

$$\Pr(Y \leq 495) \approx \Pr\left(Z \leq \frac{495 - 450}{15}\right) = \Pr(Z \leq 3).$$

Thus

$$\Pr(Y > 495) \approx 1 - \Phi(3) = 1 - 0.9987 = 0.0013.$$



Continuity note

This course example uses the direct CLT setup. A continuity correction would be a later refinement.

Law of Large Numbers: stability of averages

Convergence in probability

A sequence $Z_1, Z_2, \ldots$ converges to b in probability if, for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \Pr(|Z_n - b| < \varepsilon) = 1.$$

We write

$$Z_n \xrightarrow{p} b.$$

Law of Large Numbers

If $X_1, \ldots, X_n$ are i.i.d. with finite mean μ, then

$$\overline{X}_n \xrightarrow{p} \mu.$$

Interpretation

The sample mean does not become constant. It becomes increasingly likely to be close to the true mean.

LLN vs. CLT vs. Chebyshev

Three tools, three purposes

Tool	What it gives	When to use it
Markov	one-sided upper bound	nonnegative variable; mean known
Chebyshev	two-sided upper bound	mean and variance known; distribution unknown
LLN	long-run stability	explaining why averages approach the mean
CLT	normal approximation	large sums or sample means; finite variance
Exact normal	exact probability	variables are normally distributed

Important distinction

Chebyshev is a guarantee; the CLT is an approximation; exact normal calculations are exact under normality.

Final checklist

Before solving

- Define the random variable and the event.
- If normal, standardize using $Z = (X - \mu)/\sigma$.
- If working with a sample mean, use $\mathrm{Var}(\overline{X}_n) = \sigma^2/n$.
- If only mean/variance are known, use Markov or Chebyshev; do not invent a distribution.
- If using the CLT, identify μ, σ^2, and whether you are approximating a sum or a mean.
- For two-sided normal probabilities, convert to $2\Phi(c) - 1$ when symmetric.

Main message

Most errors come from using the right formula on the wrong random variable. Set up the random variable first.