Lesson 50
Experiments and Probability

NAME:

Start by navigating to the Online Lesson for instructions.

Objectives

- Compare two treatments in an experiment to decide if the differences are significant.
- Analyze decisions using inferential statistics.

Why?

Experiments are important because they help develop new products, treatments, and discoveries. This occurs in many fields, including medicine, technology, farming, sports, and more. Researchers need to remain objective when conducting experiments to avoid introducing bias.

Warm Up

1) In your own words, define experiment.

2) If the z-score is 1.960, what is the confidence level? How does this relate to standard deviations?

To continue, return to the Online Lesson.

Explore

Analyzing Experiments

Fill in the notes as you watch the video in the Online Lesson.

Experiments

- An experiment is a type of study in which a researcher:
  - Establishes _____ (experimental) groups
  - Designs a plan to _____ _____ that intentionally changes a behavior

- The treatment group receives the treatment or is _____ _____.
- The control group does not receive the treatment, or _____ _____ if the control and treatment groups are anonymous.
- In inferential statistics, an experiment must be _____ _____ to determine if it is successful.
- Through analysis, researchers:
  - _____ the treatment and control groups
  - Draw _____ and make _____ about the group(s)
  - Apply conclusions and inferences _____

Simulation

- Because it is not practical to repeat live experiments over and over, _____ are used to essentially shuffle, reorganize, and reassign data, ensuring random chance.
- A randomized simulation asks this question (about your live experiment results): _____ _____
- The question is answered using complex simulation technology to:
  - _____ all of the data together
  - Randomly re-assign (sort, shuffle) the individuals into the same-sized treatment and control groups as the _____
  - _____ observed differences between the groups
  - Run the simulation _____ times to get a reliable _____ _____

Scenario-Based Tasks

- A _____ is a real-life situation or story requiring further analysis.
- You will analyze the scenario to _____ and/or make recommendations.
- Because the tasks often have more than one answer, you must _____ your reasoning and _____ your decisions.
- For scenario-based tasks:

1) _____

- The observed difference:

$$\text{observed difference} = |\text{treatment mean} - \text{control mean}|$$

- The z-score for the observed difference:

$$z = \frac{\text{observed difference}}{\text{simulated standard deviation}}$$

2) _____ the z-score by answering, "What does this mean?"

z-Value	Significance Level	Meaning
$\lvert z \rvert < 1.96$	Not significant	Could occur by chance
$\lvert z \rvert > 1.96$	5% (a.k.a. 95% certainty)	Statistically significant

Remember, the percentages are not used in the calculations, but they help you see if the results are unusual (significant).

3) _____ (so now what?)

- _____ significance asks: "Is this real or due to chance?"
- _____ significance asks: "Does this matter?"
- _____ asks: "Is the cost worth it?"
- _____ asks: "What else, if anything, matters?"
- Justify your thinking using data in a _____.

Example 1

Complete the example as you watch the video in the Online Lesson.

Medication Trial

A pharmaceutical lab tested a new sleep medication against a placebo. Forty volunteers with insomnia (trouble sleeping) were randomly assigned to either the medication or the placebo group. Each participant recorded the number of hours they slept each night for a week. The observed difference between the two groups was 1.40 hours.

Group $n = 20$	Average Hours of Sleep
Medication	7.33
Placebo	5.93

After the experiment, technology simulated $10{,}000$ trials, resulting in a standard deviation of 0.37 hours.

A) Calculate the z-score for the observed difference. Explain if you think the medication is effective.

B) If the medication was not effective, what would you expect the results to be?

C) Headline: "*Amazing Sleeping Pill Guarantees 8 Hours of Rest! Study Confirms medication increases sleep 140%!*" Is the percentage in the headline accurate or misleading? Explain.

D) What is not mentioned in the headline that should be reported?

If needed, continue your work on a separate sheet of paper.

Example 2

Complete the example as you watch the video in the Online Lesson.

Grocery Store Checkout

The Food-Mart grocery store chain wanted to know if it should add a new self-checkout system to its stores. During peak hours, 600 customers were randomly assigned to self-checkout or a cashier. Then, simulation technology was used to complete $10{,}000$ randomized tests for checkout time based on the results of the experiment. The standard deviation of the checkout time was 0.34 minutes, and the error rate standard deviation was 2.78 percentage points.

Group $n = 300$	**Checkout Time (minutes)**	**Error Rate**
Treatment: self-checkout	3.2	12%
Control: cashier	4.1	3%

Business considerations:

- One self-checkout system costs $\$45{,}000$.
- Cashiers earn $\$16$ per hour and handle an average of 15 customers per hour.
- Each error at self-checkout requires 5 minutes of staff time to resolve at a rate of $\$20$ per hour.
- The stores average $2{,}000$ customers daily.

A) Calculate the checkout time z-scores for the observed difference. Explain the significance at a 5% level.

B) Calculate the error z-score for the observed difference. Explain the significance.

C) What other factors beyond statistics and cost should be considered?

If needed, continue your work on a separate sheet of paper.

Example 3

Complete the example as you watch the video in the Online Lesson.

Farm Fertilizer

A farming cooperative selected 80 plots of the same size with the same conditions to test a new organic fertilizer. The results were run through a randomized simulator $10{,}000$ times, resulting in a standard deviation of the simulated differences of 4.15 bushels. Cost considerations: Organic fertilizer costs an additional $\$45$ per acre compared to standard fertilizer. The current market price is $\$6.50$ per bushel.

Group $n = 40$	**Mean Yield in Bushels per Acre**
Treatment: new organic fertilizer	168.3
Control: standard fertilizer	162.7

A) Calculate the z-score for the observed difference. Explain the statistical significance at the 5% level.

B) Calculate the expected financial benefit per acre for a yield increase. Would switching to the organic fertilizer be profitable?

C) As a decision maker for the farm co-op, what recommendations would you make? (Consider statistical and economic factors.)

If needed, continue your work on a separate sheet of paper.

Checkpoint: Analyzing Experiments

Using the scenario from Example 3, to avoid losing money, what is the minimum market price per bushel required to be able to afford the new fertilizer?

To continue, return to the Online Lesson.

Practice 1

Complete problems on a separate sheet of paper.

Use the following scenario for problems 1–5.

A small restaurant group tests two menu designs to see if they impact the amount of money diners spend on their meals. They randomly selected 200 customers. By table, customers were either provided a new menu with photos and text or the standard menu with text only.

Group $n = 100$	Average Check Amount
Treatment: new menu	\$31.95
Control: standard menu	\$28.65

After the experiment, a simulation test with 10,000 trials was conducted. The standard deviation of the simulated differences was \$1.66. An observed difference of \$3.30 occurred in 234 simulations.

Implementation details: New menus cost \$7.50 each and are replaced every six months (compared to replacing the old menus every 2 years). An average of 6,000 customers are served monthly across all four locations.

1) Calculate the z-score for the observed difference. Explain the statistical significance at the 5% level.

2) Calculate the expected additional yearly revenue (earnings) if the new menu were to be implemented.

3) Determine the annual cost for implementing 200 new menus at all locations.

4) Based on the revenue and costs, should the restaurant group adopt the new menu? Explain.

5) What assumptions did you make to complete the cost analysis? What additional information would strengthen your recommendation?

Use the following scenario for problems 6–8.

A technology company claimed its brain games increased test scores for college students. The company randomly selected 400 students to participate. The games cost \$3.75 per month, billed annually. A 10,000 trial simulation was performed assuming no difference between treatments.

Group $n = 200$	Mean Test Score
Treatment: brain games	81.5% (B–)
Control: placebo games	79.8% (C+)
Standard deviation	5.86

6) Calculate the z-score. Explain the statistical significance at the 5% level.

7) Press Release: "Brain games improve students' test scores by a full letter grade! Sign up now to improve grades for a minimal cost." Does the technology company have evidence to support this headline? Explain.

8) Suppose you purchased a month of the brain games and had no change in your test scores. Write a short response requesting a full refund. Remember to use statistics to justify your reasoning.

Use the following scenario for problems 9–12.

A physical therapy office was testing a new rehabilitation technique on 90 patients recovering from shoulder surgery.

The new technique requires \$15,000 in new equipment and a \$2,000 certification course for physical therapists.

Group $n = 45$	Recovery in Weeks	Pain Score (0–10)
Control: standard technique	13.8	3.6
Treatment: new technique	11.2	2.8

The simulation results showed a simulated difference standard deviation of 1.12 weeks of recovery and 0.62 points for pain scores.

9) Calculate the z-score for the observed differences of both recovery and pain score. Explain the statistical significance at the 5% level.

10) From the patient's perspective, evaluate the practical significance of both recovery time and pain score.

11) Suppose the office treats 200 shoulder patients annually, and three physical therapists would need to be certified if the new technique were implemented. Should the practice invest in the new technique? (Remember to consider statistics as well as practical outcomes.)

12) Explain what might have resulted in the variables showing mixed results. What could this mean for the physical therapist when making decisions?

Use the following scenario for problems 13–15.

A high school statistics class conducted an experiment to test whether using phones before bed affects sleep quality. The class randomly selected 150 schoolmates to participate.

The class used a randomization test to simulate $10{,}000$ trials. The standard deviation of the simulated differences was 0.41 hours.

Group $n = 75$	**Mean Hours of Sleep**
Treatment: no phone one hour before bed	7.7
Control: normal phone use	6.8

Additional data: Health experts recommend teenagers sleep 8–10 hours each night. Student participants who slept more reported feeling more alert and scored an average of four points higher on weekly tests.

13) Calculate the z-score for the observed differences. Explain the statistical significance.

14) Is the treatment meaningful despite it falling short of the recommended 8–10 hours of sleep each night? Explain.

15) Would you personally try putting your phone away one hour before bed? Explain your decision and consider the challenges of following this rule.

To continue, return to the Online Lesson.

Mastery Check

Show What You Know

The track coaches at the Colossal Speed Camp tested a new warm-up method for 80 teen athletes running the 100-meter sprint.

Group $n = 40$	Mean Sprint Time	Mean Pulse-Rate Recovery Time
Treatment: dynamic stretching	13.8 sec	4.2 min
Control: static stretching	14.3 sec	4.7 min
Observed difference	0.5 sec	0.5 min
Standard deviation	0.21 sec	0.38 min

Simulation Results: A randomized test with $10{,}000$ simulations was performed. The observed difference and standard deviation are reported in the table.

Additional Information:

- Dynamic stretching: stretching while moving
- Static stretching: stretching while holding
- Pulse-rate recovery: a measure of cardiovascular fitness, faster recovery means an athlete is more in shape and has a lower risk of injury
- In competitive races, 0.5 seconds can be the difference between first and last place.
- Coaches expect athletes to compete in races from November through June.

A) Calculate both z-scores. Explain the statistical significance at the 5% level.

B) Practically speaking, why do you think the track coaches also monitor pulse-rate recovery time?

C) What type of stretching would you recommend the track coaches require for athletes during the season? Consider the statistical and practical implications for athletes.

Say What You Know

In your own words, talk about what you have learned using the objectives for this lesson and your work on this page.

To continue, return to the Online Lesson.

Practice 2

Complete problems on a separate sheet of paper.

Use the following scenario for problems 1–3.

A software company tested a system update claiming to improve battery life. They randomly select 250 users to participate in an experiment.

A randomized simulation with 10,000 trials resulted in an observed difference of 0.3 hours. The standard deviation of the simulated differences was 2.73 hours.

Group $n = 125$	Average Battery Life
Control: current software	12.9 hours
Treatment: new software	13.2 hours

Marketing claim: "Our revolutionary new software increases battery life by an average 18 minutes, allowing for more productive screen time every day."

1) Calculate the z-score. Explain the statistical significance at the 5% level.

2) Explain any concerns you have about the marketing claims.

3) Software updates are important for security, performance, and system stability. Practically speaking, should software be updated even if the battery life is not significantly improved?

Use the following scenario for problems 4–6.

A pharmaceutical research group tested a new pain relief medication. Researchers randomly assigned 100 patients with chronic pain to one of two groups.

A randomized test with 10,000 simulations was performed. The standard deviation of the simulated differences was 0.76 points.

Group $n = 50$	Mean Pain Scale Score (0–10)
Treatment: new	3.8
Control: placebo	2.1

Pain score: A patient pain score of zero means there is no pain. A patient pain score of ten is the highest level of pain.

4) Calculate z-scores for the observed difference. Does the new medication have statistical significance at the 5% level?

5) If you were looking at the research for the pharmaceutical company, would you recommend releasing the medication to a larger group? Justify your response.

6) What assumptions about the experiment and participants should be verified before making any final recommendations?

Use the following scenario for problems 7–10.

The environmental science club tested which type of cup kept coffee hot the longest. They randomly assigned 60 cups of coffee (all starting at 180°F) to two groups. The club asked the statistics students to simulate 10,000 trials and record the standard deviation of the differences.

Group $n = 30$	Temp after 30 min	Still Drinkable at 45 min
Treatment: insulated travel cup	156°F	86%
Control: disposable paper cup	128°F	43%
Standard deviation	7.8°F	11.2%

The environmental science club notes:

- The cost of the coffee was not included in this analysis.
- The most popular reusable travel cup used at school costs $18.
- Disposable paper cups cost $0.15 each.
- Students who use disposable cups report that they use them so they don't have to worry about losing a cup.
- A student would use 180 disposable cups if they used one per day for an entire school year.

7) Calculate both z-scores for the observed difference. Explain the statistical significance at the 5% level.

8) The environmental science club wants to show that reusable cups are more economically as well as environmentally friendly. How many weeks would it take for the reusable cup to pay for itself?

9) Club poster: "Purchase your reusable travel cup today for only $18! Your coffee stays warmer longer, tastes better, AND is significantly better for the environment." Does the experiment support the club poster? Explain.

10) Why do you think it is useful to use two variables (temperature and drinkability) rather than temperature alone?

Use the following scenario for problems 11–13.

Driver's education instructors tested whether listening to music affects the reaction time of teen drivers in a virtual reality simulator. They randomly assigned 90 teen drivers to two groups. After the experiment, a randomization test was run $10{,}000$ times. The observed difference was 0.07 seconds, and the standard deviation was 0.053 seconds.

Group $n = 45$	Mean Reaction Time
Treatment: music	0.78 sec
Control: no music	0.71 sec

Highway safety report: At 60 miles per hour, a car travels 88 feet per second. Even a 0.1 second delayed reaction can be dangerous in certain driving conditions.

11) Calculate the z-score for the observed difference. Is the difference in the reaction time statistically significant?

12) Practically speaking, why might the reaction time still be important for driving safety, especially for teen drivers?

13) Based on your analysis, should driver's education instructors recommend teen drivers avoid listening to music?

Use the following scenario for problems 14–16.

A psychology teacher randomly assigned 84 of her students to different note-taking methods during class, and then tested their memory. A randomized simulation run $10{,}000$ times resulted in a standard deviation of 1.92 percentage points.

Group $n = 42$	Mean Test Score
Treatment: handwritten notes	84.6%
Control: typed notes (laptop)	77.4%

Class Observations:

- Students typing their notes averaged 350 words per class.
- Students writing their notes by hand averaged 180 words per class.
- Laptop users admitted to checking their email or going online at least once during class.

14) Calculate the z-score for the observed difference. Explain the statistical significance at the 5% level.

15) Considering what students wrote about distractions, what might explain writing nearly twice as much but receiving a lower test grade?

16) Suppose you are a student tutor working with a student who only types their notes and is struggling with their tests. What recommendations would you make to the struggling student?

To continue, return to the Online Lesson.