

Lesson #57 · Topic 4E · Practice

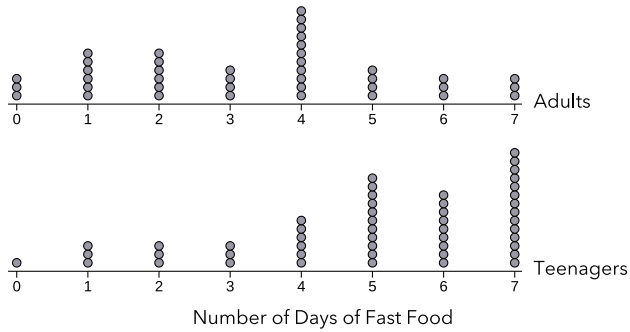
Name _____

1. The inference procedure for estimating a difference of means is called a “two-sample t -interval for $\mu_1 - \mu_2$ ”.
 - a. Why does the name of the procedure include “two-sample”?
 - b. Why does the name of the procedure include “ t ”?
 - c. Why does the name of the procedure include “interval”?
 - d. Why does the name of the procedure include “for $\mu_1 - \mu_2$ ”?

2. A recent experiment done at the University of Minnesota Law School investigated the impact that AI could have on a lawyer's ability to perform certain tasks. The 59 volunteers for the experiment were randomly assigned to two groups: 30 of the volunteers were asked to write a draft of a client memo with the help of ChatGPT, while the remaining 29 volunteers were asked to write a draft of a client memo without any AI assistance. For the group that used ChatGPT, the average amount of time it took to complete the task was 215.7 minutes with a standard deviation of 73.0 minutes. For the control group, the average amount of time it took to complete the task was 244.4 minutes with a standard deviation of 58.0 minutes. The distributions of completion time were fairly symmetric with no outliers for each group. The researchers want to construct and interpret a 95% confidence interval for the true difference in mean times it takes to write a draft of a client memo (ChatGPT – control).
- Check the conditions. For each condition that is met, explain why it is important to check the condition.
 - Using the conservative degrees of freedom (using the lower of the two group sizes), find the critical value t^* for a 95% confidence interval.
 - What is the degrees of freedom given by technology? Use it to find the critical value t^* for a 95% confidence interval.
 - Use technology to calculate the confidence interval. Do the researchers have convincing statistical evidence that using AI decreases the mean amount of time needed to draft a client memo? Explain.
 - If the conservative degrees of freedom had been used to calculate the interval, would it be wider or narrower than the interval in part (d)? Explain.

3. Reese plans to move to a big city after college and is doing some research on rent prices for one-bedroom apartments. From a random sample of 40 apartments available in New York City, she calculates a mean of \$2,238 and a standard deviation of \$567. From a separate random sample of 50 apartments available in Chicago, she calculates a mean of \$1,968 and a standard deviation of \$433.
- a. Construct and interpret a 95% confidence interval for the true difference in mean rent prices for all one-bedroom apartments available in New York City and Chicago.
- b. Based on the confidence interval, does Reese have convincing statistical evidence that the mean rent price for one-bedroom apartments is greater in New York City than in Chicago?

4. Adeola selected independent random samples of 40 adults and 50 teenagers and asked them to report the number of days in the past week in which they ate fast food. Here are graphs of the results:



Is the sampling distribution for the difference in sample means approximately normal?

- A) No, because the population standard deviations are unknown.
- B) No, because the distribution of responses from teenagers is strongly skewed left.
- C) Yes, because of the central limit theorem.
- D) Yes, because it is safe to assume that 40 is less than 10% of the population of all adults and 50 is less than 10% of all teenagers.
5. Independent random samples of 30 basketball players were selected from the women's National Basketball Association and the men's National Basketball Association. Based on these samples, the women have a mean annual salary that is \$7.92 million less than the men. A 99% confidence interval for the true difference in mean annual salary (women – men), in millions of dollars, is $(-8.27, -7.57)$. Suppose that this process was repeated many, many times. Which of the following statements is true?
- A) About 99% of the intervals will be below -8.27 .
- B) About 99% of all the intervals created will capture the difference of population means.
- C) About 99% of the intervals will have values between -8.27 and -7.57 .
- D) About 99% of all the intervals created will capture the true difference of sample means of $-\$7.92$ million.