Lab 11b: C.I. for Sample Means

Stat 131A, Fall 2018

Learning Objectives:

- How to use a sample mean to estimate a population mean.
- Computing confidence intervals for population proportions.

General Instructions

- Write your solutions in an `Rmd` (R markdown) file.
- Name this file as `lab11b-first-last.Rmd`, where `first` and `last` are your first and last names (e.g. `lab11b-gaston-sanchez.Rmd`).
- Knit your `Rmd` file as an html document (default option).
- Submit your `Rmd` and `html` files to bCourses, in the corresponding lab assignment.

Problem 1

Weight of a rock: In a geology course, students are learning to use a balance scale to accurately weigh rocks. One student plans to weigh a rock 20 times and then calculate the average of the 20 measurements to estimate her rock's true weight. A second student plans to weigh a rock 5 times and calculate the average of the 5 measurements to estimate his rock's true weight.

Which student is more likely to come the closest to the true weight of the rock he or she is weighing?

a. The student who weighed the rock 20 times.

b. The student who weighed the rock 5 times.

c. Both averages would be equally close to the true weight.

Problem 2

A medical researcher wants to estimate the mean weight of babies born to women over the age of 40. The researcher chooses a random sample of 100 pregnant women who are over 40. Using the mean birth weight of the 100 babies in the sample, the researcher calculates the 95% confidence interval for the population mean. With 95% confidence, the researcher estimates the mean birth weight of all babies born to women who are over the age of 40 to be

between 2935 and 3135 grams. The researcher wants to maintain the 95% level of confidence but report a confidence interval with a smaller margin of error. What should she do?

a. Redo the study, but this time with a sample size 225.

b. Redo the study and choose a different sample of size 100.

c. Redo the study, but this time with a sample size 64.

Problem 3

For each of the following situations, decide if it is valid or invalid to use a confidence interval to estimate the population mean.

Situation A: In order to estimate μ, the mean annual salary of high-school teachers in a certain state, a statistics student selected a random sample of 150 teachers from the state. The student found the mean and standard deviation of the teachers' salary.

Situation B: A medical researcher wanted to estimate the mean recovery time from open-heart surgery for males between the ages of 50 and 60. The researcher followed a random sample of 15 male patients in this age group who underwent open-heart surgery. The mean recovery time of the 15 patients was 26 days.

Situation C: A statistics class wanted to estimate the height of all male students at their college. They chose a sample of 35 male students playing basketball in the gym one afternoon. The student then found the mean and standard deviation of the heights in the sample.

Problem 4

Length of a rod: Engineers on the Bay Bridge are measuring tower rods to find out if any rods have been corroded from salt water. There are rods on the east and west sides of the bridge span. One engineer plans to measure the length of an eastern rod 25 times and then calculate the average of the 25 measurements to estimate the true length of the eastern rod. A different engineer plans to measure the length of a western rod 20 times and then calculate the average of the 20 measurements to estimate the true length of the western rod.

The engineer who collected 20 measurements for the western rod uses the mean to calculate the 90% confidence interval for rod length (in feet). Her interval is (23.76, 24.23). What does a 90% confidence interval tell you about rod length in this case?

a. The engineer is 90% confident that the true length of the rod is between 23.76 feet and 24.23 feet.

b. The engineer is 90% confident that most rods of this type have lengths between 23.76 feet and 24.23 feet.

c. The engineer is 90% confident that this interval includes the mean of the 20 length measurements she took.

Problem 5

Length of a rod (follows from the previous problem). One engineer plans to measure the length of an eastern rod 25 times and then calculate the average of the 25 measurements to estimate the true length of the eastern rod. A different engineer plans to measure the length of a western rod 20 times and then calculate the average of the 20 measurements to estimate the true length of the western rod.

Suppose the engineers construct a 90% confidence interval for the true length of their rods. Whose interval do you expect to be more precise (narrower)?

a. Both confidence intervals would be equally precise.

b. The engineer who weighed the rod 25 times.

c. The engineer who weighed the rod 20 times.

Problem 6

A utility company serves 50,000 households. As part of a survey of customer behavior, they take a simple random sample of 750 of these households. The average number of television sets in the sample households turns out to be 1.86, and the SD is 0.80. If possible, find a 95%-confidence interval for the mean number of televisions sets in all 50,000 households. If this is not possible, explain why not.

Problem 7

As part of the survey described in the previous question, all persons age 16 and over in the 750 sample households are interviewed. This makes 1528 people. On the average, the sample people watched 5.20 hours of television the Sunday before the survey, and the SD was 4.50 hours. If possible, find a 95%-confidence interval for the mean number of hours spent watching television on that Sunday by all persons age 16 and over in the 50,000 households. If this is not possible, explain why not.

Problem 8

A real estate office wants to make a survey in a certain town, which has 50,000 households, to determine how far the head of households has to commute to work. A SRS of 1000 households is chosen, the occupants are interviewed, and it is found that on average, the heads of the households commuted 8.7 miles to work. The SD of the distances was 9 miles. (All distances are one-way; if someone isn't working, the commute distance is defined to be 0.)

a. The average commute distance of all 50,000 heads of households in the town is estimated as _____, and this estimate is likely to be off by _____ or so.

b. If possible, find a 95% condifence interval for the average commute distance of all heads of households in the town. If this isn't possible explain why not.