Lecture 13: Confidence Intervals

For each question, write the R code you would use and record the result when appropriate. Complete this in an RMarkdown document so that you will have the code/explanations for your personal records. Answer in complete sentences when asked to explain.

Part 1: Confidence Intervals from Summary Information

1. A company wants to estimate the average yearly salary of its employees. From a recent sample of 36 employees, the average salary was \$68,000. Assume the population standard deviation is known to be \$12,000.

   (a) Identify the sample mean, population standard deviation, and sample size.

   (b) Compute the standard error.

   (c) Construct the 90% confidence interval for the population mean salary.

   (d) Construct the 95% confidence interval for the population mean salary.

   (e) Which interval is wider? Explain why.

   (f) Interpret the 95% confidence interval in context.

Part 2: Confidence Intervals when σ is Known

2. We will now look at the `course_grade` variable in the `exam_grades` dataset within the `openintro` library.

   (a) Display the first 6 observations of the `exam_grades` dataset.

   (b) What does one row of the dataset appear to represent?

   (c) Is `course_grade` quantitative or categorical? Explain briefly.

3. **Step 1: Gathering the Information**

   (a) Find the sample mean of `course_grade`.

   (b) Find the sample size of `course_grade`.

   (c) Assume the population standard deviation is known to be 8 points. Compute the standard error.

4. **Step 2: Constructing the Confidence Intervals**

   (a) Construct a 90% confidence interval for the population mean course grade.

(b) Construct a 95% confidence interval for the population mean course grade.

5. **Step 3: Interpreting the Results**

   (a) Interpret the 90% confidence interval in context.

   (b) Interpret the 95% confidence interval in context.

   (c) In one sentence, explain why we use the z-distribution in this part instead of the t-distribution.

Part 3: Confidence Intervals when σ is Unknown

6. We will now look at the `bwt` variable in the `babies` dataset within the `openintro` library.

   (a) Display the first 6 observations of the `babies` dataset.

   (b) What does one row of the dataset appear to represent?

   (c) Is `bwt` quantitative or categorical? Explain briefly.

7. **Step 1: Gathering the Information**

   (a) Find the sample mean of `bwt`.

   (b) Find the sample standard deviation of `bwt`.

   (c) Find the sample size of `bwt`.

   (d) State the degrees of freedom for this confidence interval.

8. **Step 2: Constructing the Confidence Interval**

   (a) Find the critical value for a 95% confidence interval using the t-distribution.

   (b) Construct a 95% confidence interval for the population mean birth weight.

   (c) Interpret the confidence interval in context.

9. **Step 3: Checking with `t.test()`**

   (a) Use `t.test()` to verify your 95% confidence interval.

   (b) Does your answer match the output from `t.test()`?

   (c) In one sentence, explain why we use the t-distribution in this part instead of the z-distribution.

Part 4: Comparing Sample Sizes

10. **Step 1: Creating Two Samples**

    (a) Create a sample called `temp_small` by taking the first 20 non-missing values of `Temp` in the built-in `airquality` dataset.

    (b) Create a sample called `temp_large` by taking the first 80 non-missing values.

    (c) Find the sample mean, sample standard deviation, and sample size for `temp_small` and for `temp_large`.

11. **Step 2: Constructing the Confidence Intervals**

    (a) Construct a 95% confidence interval for the population mean temperature using `temp_small`. Use `t.test()` to check your answer.

    (b) Construct a 95% confidence interval for the population mean temperature using `temp_large`. Use `t.test()` to check your answer.

12. **Step 3: Comparing the Widths**

    (a) Which confidence interval is wider?

    (b) Why does the width change when the sample size increases?

    (c) Which sample gives a more precise estimate of the population mean temperature? Explain.

Part 5: Reflection

13. In one or two sentences, explain what it means to say that a 95% confidence interval captures the true population mean 95% of the time in repeated sampling.