Lab Activity: Confidence Intervals – Single Population Mean

In this lab activity, you will explore confidence intervals to estimate a single population mean.

Student Learning Outcomes

By the end of this chapter, you should be able to do the following:

- Compute confidence intervals for a single population mean in Statcato

- Interpret and examine the properties of confidence intervals

Preliminary

Read Chapter 8 Confidence Intervals in:

  Illowsky, Barbara, and Susan Dean. Collaborative Statistics. Connexions. 2 Mar. 2010 <http://cnx.org/content/col10522/1.37/>.

Make sure you understand the following **key terms** (**LR:Key Terms**):

  inferential statistics, point estimate, confidence interval, Empirical Rule, margin of error, error bound for a population mean, confidence level, alpha, Student's t distribution, degrees of freedom

Loading Data

In this activity, you will the headwidths of a population of thatch ant (source: http://www.stat.ucla.edu/data/). This dataset is available on the Statcato web site. You can use the load dataset tool in Statcato to obtain the data:

Loading an online dataset

Go to `File > Load Dataset`.   Or click the icon in the toolbar.

- Under Online Datasets, enter the web address: http://statcato.org/labs/data/thatch-ant.csv. Select "Comma-separated values" in the dataset file type drop-down menu.

- Click **Load Dataset**.

You will use the headwidth measures (mm) in column C5.

Calculating Mean and Standard Deviation

You will compute the mean and standard deviation of the population of thatch ant headwidths.

Calculating descriptive statistics

Go to `Statistics > Basic Statistics > Descriptive Statistics`.

- **Input Variables(s)**: c5

- For **Statistics**, select **Mean** and **Standard Deviation**.

- Click **OK**

The mean and standard deviation should now be displayed. Copy the descriptive statistics to **LR: Mean and Standard Deviation**. Also record the mean and standard deviation below:

* μ = _____ σ = _____

Generating Random Samples

You will generate 50 samples of size 10 from the population of thatch ant headwidths. Then you will compute the confidence intervals for each of the 50 samples.

Generating 50 random samples of size 10

Go to `Data > Generate Random Data > Sample from Column`.

- For **Sample from Column**, select **C5 Headwidth (mm)** in the drop-down menu

- In **Store Samples in** text box, enter: c10-c49

- **Number of Samples to Generate**: 10

- Click OK

This may take a minute or two. Columns C10 to C49 should now contain 50 random samples of size 10.

Calculating Confidence Intervals—Known Population Standard Deviation

Next, you will calculate the 90% confidence intervals for the 50 samples, assuming that the population standard deviation is known.   The standard deviation of the population of thatch ant headwidths is the same as that of the headwidths data in C5.

Generating CIs—known population standard deviation

Go to `Statistics > Confidence Intervals > 1-Population Mean`

- **Samples in column**: c10-c49

- **Population standard deviation**: Select **Known** and enter the population standard deviation you have obtained earlier (*)

- **Confidence level**: 0.90

- Click **OK**.

The 90% confidence intervals should now be displayed in the Log window. Copy them to **LR: Confidence Intervals—Known Population Deviation**.

Calculating Confidence Intervals—Unknown Population Standard Deviation

Now you will calculate the 90% confidence intervals for the 50 samples, assuming that the population standard deviation is unknown.

⋮TATCATO **Generating CIs—unknown population standard deviation**

Go to `Statistics > Confidence Intervals > 1-Population Mean` (or go to the corresponding item in Dialog History)

- **Samples in column**: c10-c49
- **Population standard deviation**: Select **Unknown**
- **Confidence level**: 0.90
- Click **OK**.

Copy the confidence intervals to **LR: Confidence Intervals—Unknown Population Deviation**.

Discussion

Answer the following questions in **LR: Discussion**.

1. You found the actual mean for the population of thatch ant headwidths. Count the number of confidence intervals that contain the actual population mean μ. Do this for both the listing of confidence intervals assuming the population standard deviation σ is known and known. Find the percentage of the number of confidence intervals containing μ by dividing the number by 50.

μ assumed	Number of CIs containing μ	% of CIs containing μ
Known		
Unknown		

2. What is the expected percentage of confidence intervals containing μ? Are the empirical percentages of confidence intervals containing μ close to what you have expected?

3. What would happen to the percentage of confidence intervals containing μ if you generate more confidence intervals?

4. How are the confidence intervals assuming the population standard deviation σ is known different from those assuming σ is unknown? Where do the differences come from?