Project 3 – Estimating a Population Mean or Median Using the Bootstrap Method

Bootstrap Method

If we do not have all of the data for the population of interest we can still estimate the mean or median of that population. To do so we use the **bootstrap method**.

- If you have a sample of size *n*, draw a sample of size *n* from those values using sampling with replacement.
- Compute the mean of that sample.
- Repeat the process many times. (We will repeat the process 1000 times.)
- Find the critical values for these 1000 means by finding the 2.5^{th} percentile and the 97.5^{th} percentile.
  These values separate the lower 2.5% and upper 2.5% of the means from the rest.
- The same procedure can be used for estimating a population median.

These critical values will create a **95% confidence interval** for the population mean or median. 95% of the intervals created in this fashion will contain the unknown population mean or median. We can then evaluate claims about population means or medians as plausible (it is contained by the interval) or unusual (it is outside the interval).

Using StatCrunch to Bootstrap a Population Parameter

1. Load the data into a StatCrunch spreadsheet.
2. Press Applets; highlight Resampling, then Bootstrap a statistic.
   Be sure the "From data table:" radio button is selected.
   Select the column that contains the data.
   Select "Mean" (or "Median") under the Statistic drop-down menu. Click Compute!.
3. Click 1000 times to generate 1000 Bootstrap samples with replacement and determine the mean of each sample.
4. The critical values will be labeled as 2.5th perc. and 97.5th perc.

Investigation 1

A manager of a fast food restaurant devises a new drive-through system that he believes will decrease wait time from the time an order is placed to the time the order is received. He initiates the new system at his restaurant and measures the wait time for 10 randomly selected orders. The wait times, in seconds, are provided below.

108.5	67.4	58.0	75.9	65.1
80.4	95.5	86.3	70.9	72.0

(a) Use the bootstrap method to create a 95% confidence interval for the mean wait time for the new system.

(b) The manufacturer of the system claims that the mean wait time for all customers should be approximately 80 seconds. Is this value contained inside the 95% interval?

(c) Is the manufacturer's claim plausible or is it unlikely to be true? Explain your decision.

Investigation 2

The following data represent the selling price (in thousands of dollars) of oceanfront condos in Daytona Beach Shores, Florida.

476	525	410	145	1250	360
205	425	200	205	344	1080

(a) Use the bootstrap method to create a 95% confidence interval for the mean selling price.

(b) A real estate agent claims that the mean selling price for all oceanfront condos in Daytona Beach Shores, Florida is higher than $700,000. Is the real estate agent's claim plausible or is it unlikely to be true? Explain your decision.

Investigation 3

(a) Make a claim of your own about the mean age of all COS students, and a second claim about the median age of all COS students. (The two claims can be the same if you wish, but the median is likely to be higher as the ages that are outliers are likely to be higher.)

- The mean age of all COS students is _____ years old.

- The median age of all COS students is _____ years old.

(b) Open the Project 3 data set from the StatCrunch group for this class. You will find the ages of each student in Mr. Woodbury's classes this semester. Use StatCrunch's bootstrap applet to create a 95% confidence interval for the mean and a 95% confidence interval for the median.

- Bootstrap 95% confidence interval for the mean age of all COS students:

- Bootstrap 95% confidence interval for the median age of all COS students:

(c) Based on the confidence interval for the mean age, is your claim about the mean age of all COS students plausible or is it unlikely to be true?

(d) Based on the confidence interval for the median age, is your claim about the median age of all COS students plausible or is it unlikely to be true?

(e) The COS information officer told me that the mean age is 24.4 years old and the median age is 21. Based on your confidence intervals, are those claims plausible or unlikely to be true?

Claim about the Mean Age: Plausible Unlikely to be true

Claim about the Median Age: Plausible Unlikely to be true