# Confidence Interval Creation Applet: Demonstration Guide

Purpose:

The interactive Confidence Interval Creation Applet allows instructors to demonstrate how sample size, alpha level, population shape, and variance affect confidence intervals.

Function:

The **population** is displayed as a histogram with blue bars (**A**). Population shape can be altered by selecting one of the pre-set shapes on the right side of the interface (**B**), or simply by clicking and dragging the cursor across the graph to create a custom population. Clicking just below the x-axis on the graph will set the bin above the cursor to 0. Population parameters are displayed above the graph (**C**), and the location of the population mean is shown as a blue line below the graph (**D**).
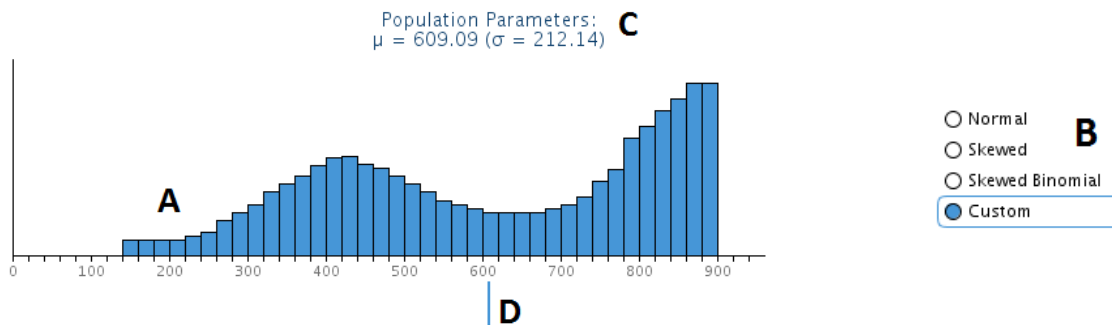
Samples can be drawn by clicking on the Sample button (**E**). The size of the next sample can be adjusted between n=6 and n=40 by sliding the yellow sample size slider (**F**) to the left or right before drawing a sample. Sample statistics (mean, s, and CI limits) are displayed in the output box (**G**). Samples are displayed as yellow bars on top of the blue population graph (**H**).
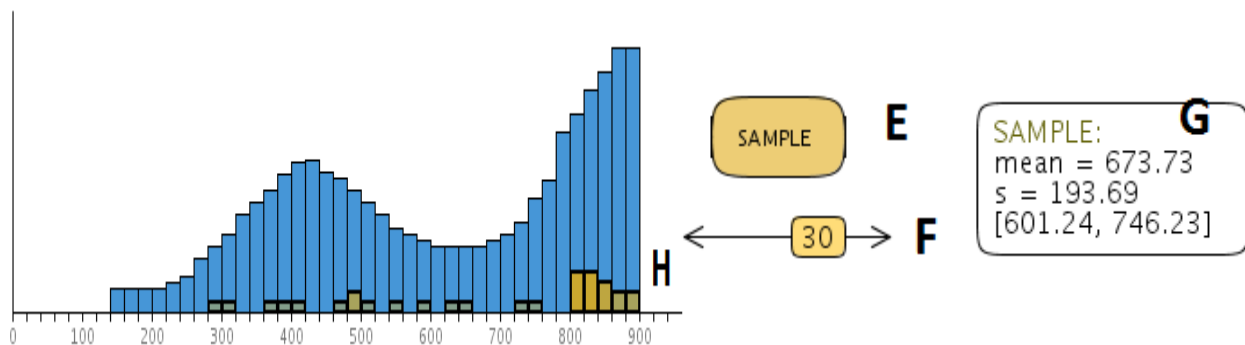
A new confidence interval is created each time a sample is drawn. The current sample's confidence interval is drawn as a horizontal bar below the population, with the sample mean as a dot in the center of the bar. Intervals are displayed in black if they contain the population mean (if they intersect with the blue population mean line), and in red if they do not. Each time a new sample is drawn, a new interval is displayed, and the last interval drops down. Up to 20 older intervals gather at the bottom of the applet. The Reset button may be used to clear the data.

Population Parameters:
μ = 500.00 (σ = 100)

485.43 ⊢——⊣ 553.24
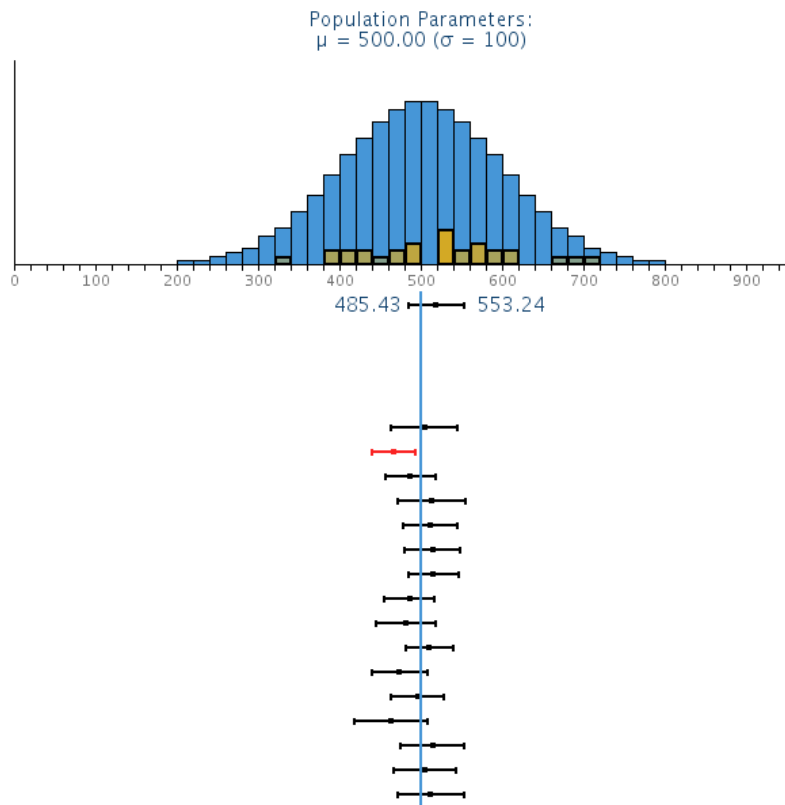
Example Cover Story

Students are probably familiar with standardized tests like the SAT which may have a mean of 500 and standard deviation of 100. We would like to know how accurate a sample mean is likely to be if we collect data from a random sample of students.

The blue distribution in the applet represents scores for the population of all students who may take the test, with all possible scores along the x-axis. We can simulate drawing a random sample of students from that population to estimate the population mean. A confidence interval can be constructed to represent the location and precision of our estimate. But does that interval contain the actual population mean?
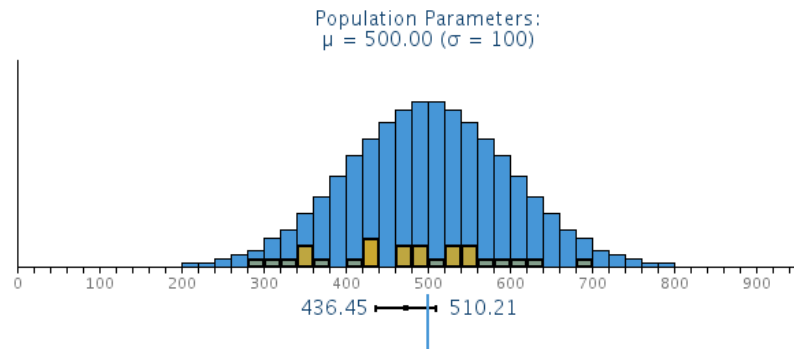
In the real world, there may be no way to know if a confidence interval contains the population mean. For demonstration purposes, this applet shows whether or not each sample produces a confidence interval that contains the population mean. If the interval is black, it contains the mean; if the interval is red, it does not. Here is where the applet facilitates a correct interpretation of a confidence interval. The applet allows the user to draw 20 samples quickly, and display a confidence interval for each one. For 95% confidence intervals, only about one of every 20 intervals calculated this way will not contain the population mean if assumptions are met. Thus, students can see that it is not the case that the population mean falls within a given confidence interval 95% of the time, but that 95% of intervals constructed in this way capture the population mean (if all statistical assumptions have been met).

The sample size and confidence level can be manipulated to show how intervals are affected by these factors. A larger sample size is expected to yield tighter intervals because the standard error of the mean decreases as sample size increases. Similarly, lower confidence levels (90% compared to 99%) also provide smaller intervals. However, 90% confidence intervals are expected to fail to include the population mean 10% of the time.

It is important for students to appreciate the variability in sampling distributions and understand that in practice researchers typically have data from only one sample, and thus they have only one confidence interval. That confidence interval may or may not include the population mean.

The instructor can change the sample size or the confidence level to demonstrate how confidence intervals are likely to be affected. The applet also allows the instructor to demonstrate how non-normal population shapes (such as skewed distributions) can result in the failure of confidence intervals. The instructor can click and drag across the distribution to create any desired shape.

It is important for students to understand that in practice, distributions may not be normal. If the population distribution of test scores was strangely skewed like the example below, 95% confidence intervals may not be as likely to contain the population mean as they should. In these instances resampling methods such as bootstrapping may be more useful than calculating a standard confidence interval using t-scores.

Population Parameters:
μ = 258.08 (σ = 230.87)