WEEK 6

DO BOOTSTRAP CONFIDENCE INTERVALS WORK? By Chris Wild

In the last video, we gave the basic idea of how to construct a bootstrap confidence interval to capture the true value.

We have an estimate calculated from our data. It might be a mean, a median, a percentage, or something else, but we know it's not going to be quite right.

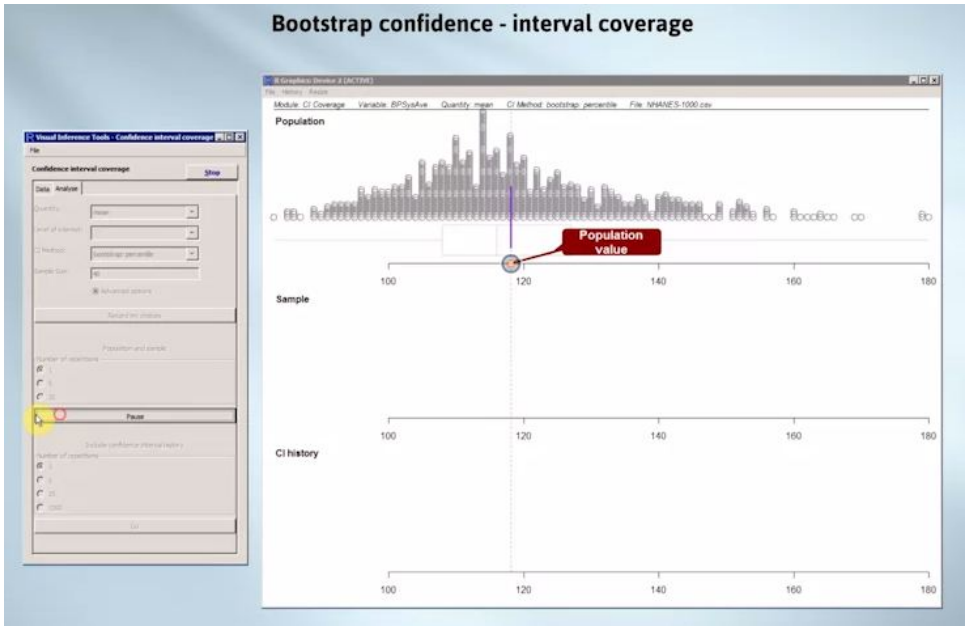
We calculate the same quantity for a large number of bootstrap resamples (from the data) and take the minimum to the maximum of what we've seen as the interval of values we think will probably cover the truth.

That's it except for one small detail. (We trim off the extreme biggest and smallest values before reading off the interval.)

Well that's the recipe, but does it work? How can we know whether a method works?

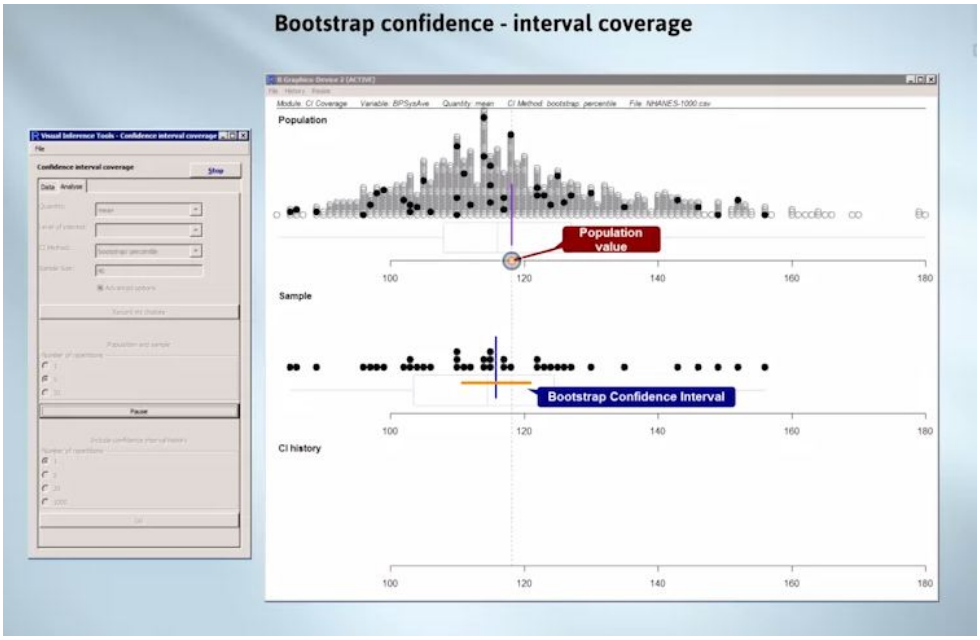
We try it out in a large number of computer generated situations where we have the population and thus know the truth, something that'll never happen in practical applications. This process is called computer simulation.

We'll look at an example now. Along the way we'll learn more about what a confidence interval is. Again, we'll treat the NHANES-1000 people as a population to sample from, and we'll look at mean blood pressures.



Still frame from animation

Here are the blood pressures for everyone in our population. We've marked the position of the mean for this population. That's the truth, that's the target we want to hit with the intervals we get from the samples we take.

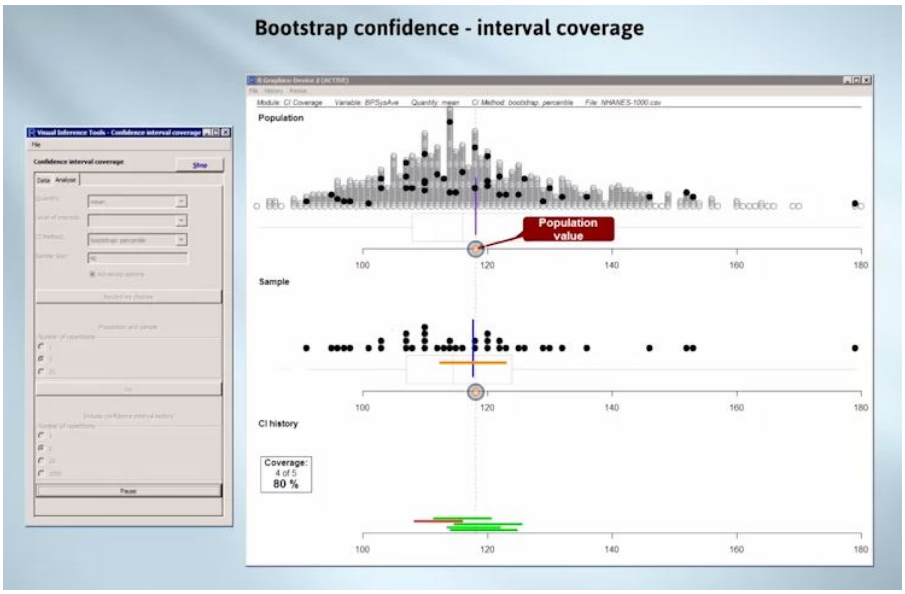


Still frame from animation

Here's the first sample being selected, falling down into the middle panel. It's off calculating the bootstrap confidence interval. There's the interval. The target value's in the interval. In statistical language our interval has covered the true value.

We'll do this a few more times. New sample, new interval. Success, we've covered the truth again. New sample, new interval, got it again.

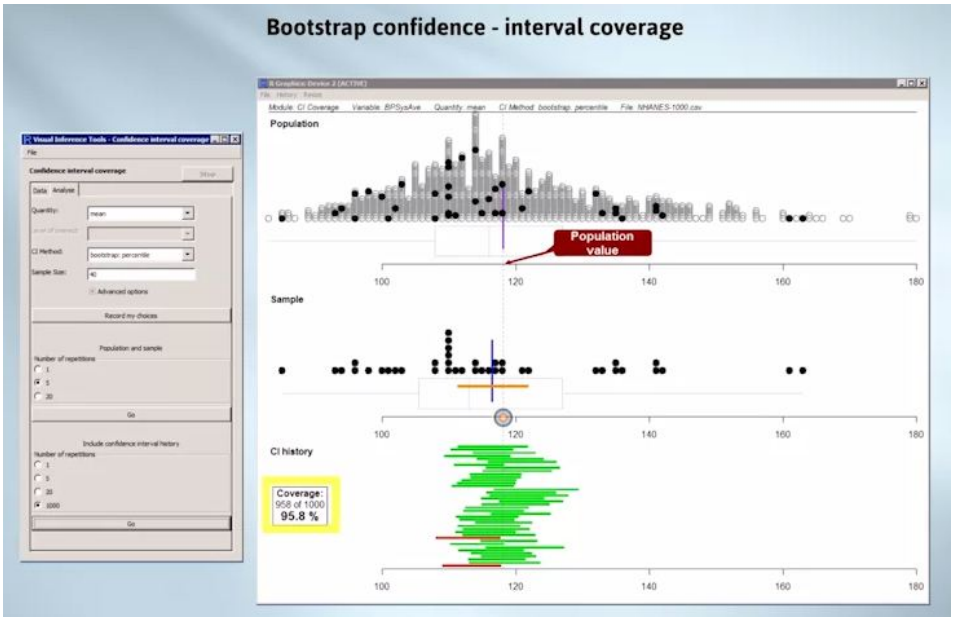
Every time we took a sample, our interval contained the true value, but it doesn't always happen.



Still frame from animation

We'll start keeping a record of the intervals we've seen and what happened. We'll drop each interval after it forms into the bottom panel. We'll colour them green if they catch the true value. Here's five more. These two have both caught the truth, so they're all green. This one missed. The interval didn't include our target (the population mean) and so we've coloured it red. Here I want to draw your attention to the coverage statistics. We've covered the truth five of the six times we've tried this. That's a success rate of 83%.

Now let's try it 1,000 times.



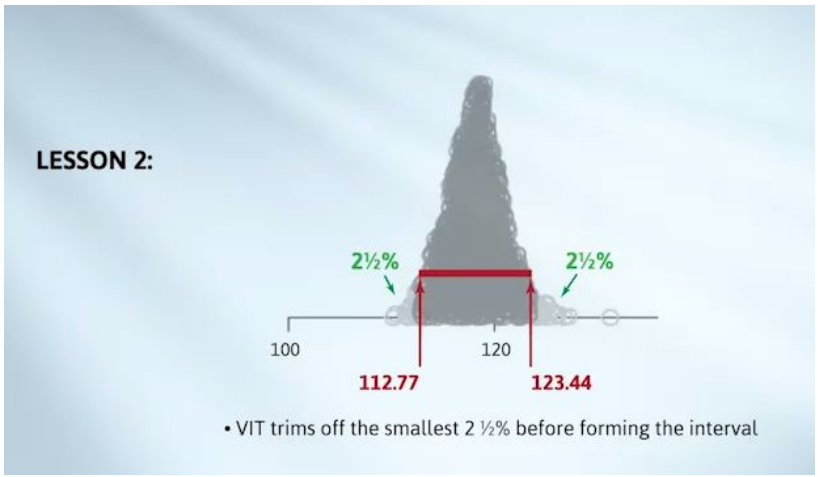
Still frame from animation

Mostly they've worked, but there's a scattering of red lines where we've failed. Overall we have a success rate that was almost 96%. For this situation, the bootstrap method of constructing an interval has almost always worked. The intervals almost always capture the true value.

Statisticians have investigated how well bootstrap confidence intervals work using simulations like this for a large variety of quantities (means, medians, percentages, et cetera) in many types of situation. They work for almost all quantities provided the sample sizes are not too small. Statisticians have also devised sophisticated modifications to make them work much better in small samples, but that's beyond this course.

LESSON ONE
So take-home lesson one from this discussion is that bootstrap resampling is a great general purpose way of constructing confidence intervals.

LESSON TWO
Lesson two concerns properties of confidence intervals themselves, but it needs a little more development.



VIT's bootstrap confidence interval construction module trims the smallest 2 1/2% and the largest 2 1/2% of the resample values before forming the interval.

When we do it this way (with large enough samples) the intervals produced by 95% of samples capture the truth. 5% miss (often not by much).

Lesson two is the defining property of a 95% confidence interval. An interval is a 95% confidence interval if it's been produced using a method that covers the true value for 95% of samples taken. But in a real application, we have one sample, and one interval. Is it one of the 95% that caught the truth, or was it one of the 5% that missed?

We never know. Our confidence in our interval comes from the fact that the method we are using almost always works. That brings us to the end of this video. In the next video, we'll go on to apply confidence intervals in data analysis.