

Textbook Exercises

3.102, 3.104, 3.105, 3.116, 4.17, 4.18, 4.21–4.25, 4.30, 4.32, 4.36, 4.40, 4.52, 4.60, 4.62, 4.63

Computer Exercises

For each R problem, turn in answers to questions with the written portion of the homework. Send the R code for the problem to Katherine Goode. The answers to questions in the written part should be well written, clear, and organized. The R code should be commented and well formatted.

R problem 1 Consider again the data in 3.116. Using 1,000,000 bootstrap replicates, use the `boot()` function from the `boot` library to find 95% bootstrap confidence intervals for the population mean. Here is sample code to do this.

```
library(boot)
foo = c(8,10,7,12,13,8,10,50)
my.mean = function(x,indices) {
  return( mean(x[indices]) )
}
boot.out = boot(foo,my.mean,1000000)
boot.ci(boot.out)
```

This will take a few seconds to calculate.

1. Describe the differences among the four intervals. Where are they centered? How wide are they?
2. The variable `boot.out$t` contains the sampled 1,000,000 means. Check to see that the “normal” bootstrap interval agrees with the mean plus or minus 1.96 times the estimated standard error. (Note the use of 1.96 instead of 2, which comes from a more precise calculation of the quantiles for the middle 95% of a perfect normal distribution.)
3. Check to see that the percentile bootstrap matches what you find using `quantile()` on the sampled means.
4. Make a density plot of the sampled means. Describe the appearance of the bootstrap distribution. Add to this plot an indication of the endpoints of each confidence interval.
5. How does the BCa confidence interval differ most from the others?

R problem 2 The data set **CaffeineTaps** contains two samples of size 10 where each value is the number of times a student tapped his finger in a minute. One group had consumed 200mg of caffeine in coffee two hours before the experiment, and one group had decaf. Details of the experiment are on page 240.

1. Use the bootstrap to estimate the mean number of taps per minute with a 95% confidence interval separately for each group. Use both the SE and percentile versions.
2. Use the bootstrap to estimate the difference, caffeine minus decaf, in the population means, with a 95% confidence interval. Use both the SE and percentile versions.
3. Compare the widths of the confidence intervals for the individual population means and for the difference in population means. What is the ratio of these widths? Which is larger?
4. Write an interpretation of one of the confidence intervals for the difference in means.