

Assignment #5
Due Thursday, February 14
R and the Bootstrap

Write an R function that takes as input a vector v and a number th , and outputs the fraction of the entries in the vector that are greater than or equal to th . For example, if v equals $c(3,1,5,8)$ and th equals 4 the function should output 0.5. Try to avoid loops to make the program run faster.

Write a bootstrap function to calculate confidence intervals.

- a. This function inputs a vector, a confidence interval width (for example 0.95 or 0.75), the number of times to take a bootstrap sample, and a function that will calculate the quantity of interest (as in lecture use `...` for any arguments to this function).
- b. Use the built-in R function “sample” to take bootstrap samples (in R type, “`help(sample)`” to learn more about this function).
- c. Each bootstrap sample should be “with replacement” and the same size as the original vector.
- d. For each bootstrap sample calculate the quantity of interest. When you have done this for the specified number of bootstrap samples, you will have a vector. Use this vector to calculate the confidence interval. The built-in R function “sort” will be helpful (in R type “`help(sort)`” to learn more about this function).

In Content I have placed a file “DataforHW5.csv”. Use the built-in R function “read.csv” (in R type, “`help(read.csv)`” to learn more about this function).

1. Calculate the fraction of numbers in column 1 that are greater than or equal to 90. Use the bootstrap to calculate the 95% confidence interval for this quantity.
2. Calculate the fraction of numbers in column 2 that are greater than or equal to 60. Use the bootstrap to calculate the 75% confidence interval for this quantity.

Turn in your code, and the answers to #1 and #2.