

Statistical Methods for Data Science

Mini Project 4

1. Write R code for computing nonparametric bootstrap confidence intervals for a population parameter based on a random sample from the population. You have an example code that computes these intervals for population median using ‘boot’ package in R. Here you are asked to *write your own code*. The code would also need to:
 - compute the following three bootstrap confidence intervals: normal approximation, basic bootstrap, and percentile bootstrap; and
 - compute bias and standard error of the parameter estimate as they are needed for the confidence intervals.
2. Apply the code in (1) to perform **inference on the 90-th population percentile** using the CPU time data (available in the file cputime.txt). Be sure to report the following: parameter estimate, estimated bias and standard error, the three 95% bootstrap confidence intervals for the parameter, and your conclusions.

2 Bonus points will be given for good, neat work.

Instructions:

- Due date: Thursday, November 3.
- Total points = 10
- Submit a typed report and include all relevant plots.
- You can work on the project either individually or in a group of no more than two students. In case of the latter, submit only one report for the group, and include a description of the contribution of each member.
- Do a good job.
- You must use the following template for your report:

Mini Project #

Name

Names of group members (if applicable)

Contributions of group members

Answers and justifications for each exercise

Provide the R codes in an appendix. Your code must be annotated. No points may be given if a brief look at the code does not tell us what it is doing.