POLI706: Advanced Methods of Political Analysis

Problem set 10

For this entire homework, you may not use the `{boot}` package in R.

Exercise 1

Summarize in your words the idea(s) of bootstrapping. You may want to consult further resources.

Exercise 2

Using some data of interest, show via R that, as the number of bootstrap draws increases, the bootstrapped coefficient estimates approach the analytically derived sampling distribution. Showing the results via one or several graphs is a rather good idea.

Exercise 3

Let's load all the packages needed for this assignment (this assumes you've already installed them).

```
#: echo: true
library(tidyverse)
library(moderndive)
library(infer)

# Example code for one-time resampling of pennies_sample
# Let's perform the resampling with replacement of the 50 slips of paper
# representing our original sample 50 pennies:

virtual_resample <- pennies_sample |>
  rep_sample_n(size = 50, replace = TRUE)
```

a. Using the sample data, `pennies_sample`, from the `{moderndive}` package, virtually resample 100 times, 1,000 times, and 10,000 times, respectively. What are the chief differences between bootstrap distributions and a sampling distribution of `pennies_sample`?

b. Show via R that, as the number of bootstrap draws increases, the bootstrapped coefficient estimates approach the analytically derived sampling distribution. Showing the results via one or several graphs is a rather good idea.

c. From the answer of Question 3a, provide 95% confidence intervals around your estimates. Prudent that you are, you'll do so via the non-parametric bootstrap. Write up the results, and offer an exceptionally pretty graph to show your insights. What condition about the bootstrap distribution must be met for us to be able to construct confidence intervals using the standard error method?