

CS112: Discovering Computer Science
Individual Project 4: Polling Accuracy

1 Overview

The fundamental research question addressed by this project is:

*How does the sample size of a poll
affect the accuracy of the poll?*

The purpose of this project is to investigate how polling size affects accuracy of the polling results. You should follow the project outlined as Project 5.1 in Section 5.8 of the book. The learning objectives of this project are to gain experience with Monte Carlo simulations, apply the MC simulations to a real-life situation, and develop additional python programming skills especially with randomness and decisions.

You will work on this project individually. You may not seek help or share code or ideas with other students in the class, or other people in general. You may seek help from the class professor and TAs. You should not have your personal tutor help you with this project. They can help you write and understand other similar Monte Carlo simulation programs, but you should complete this project on your own. Consider this project as a take home exam.

2 Project Suggestions

In class we follow the example of *think and plan first* and then we write code. Do not start writing code until you completely understand the full program and what part each will do. This program should not take more than two hours to complete (it should take about one hour). If you are spending a lot more time on it than this, that is an indication that you might need to do more planning and thinking before you begin writing the program.

You may want to spend an hour re-reading the book chapter. Now that we have spent a week in class working on Monte Carlo simulations, a second reading of the book will fill in a lot of missing questions and deepen your understanding. An hour spent with the book might save two spent on the programming!

3 Project Overview

The book outlines a good project. It discusses the motivation for the project and then provides a series of steps to help you build a program to answer the primary question. Read the problem statement carefully so that you fully understand the basic inquiry.

Follow the steps outlined in the book to build your program, one piece at a time:

- Part 1 has you build a basic function to take a poll once.
- Part 2 has you simulate many polls in order to find the two extremes of polling outcomes.
- Part 3 has you try different sample sizes and then plot the relationship between sample size and the polling extremes.
- Part 4 is not required. (Does error depend on the actual percentage).

You should complete the first three parts of the project working on your own. Be sure to follow the complete conventions for commenting your program well.

I want you to write your own `min` and `max` functions even though python has perfectly fine built-in functions. Each of these two functions will take a list of numbers as input and return either the largest or smallest value in the list. You can assume the list has at least one value in it (no empty lists allowed). Here is a summary of all the functions you will write (probably in this suggested order):

- `main()` which oversees the operation of your program.
- `poll()` as explained in the book
- `min()` as explained above.
- `max()` as explained above.
- `pollExtremes()` as explained in the book
- `plotResults()` as explained in the book

As you build each part, you might want to insert some temporary code in `main()` to test each function out. At the end, your `main()` will probably only call the `plotResults()` function which calls the others.

4 Submission

Submit your working python program. Be sure your program produces the graph asked for in the third part of the project. Be sure you follow proper conventions for documenting and organizing your program. Be sure you run and test your program one last time after saving it.