

CSIS 200: FINAL PROJECT

Stephen Clarke
Siena College, Loudonville, NY
(Dated: December 2, 2016)

I have some work to do regarding applying what I learned in class to independent projects, as I struggled with this. For the first question, I found the percentage of it raining only once in a month when there's a 20 percent chance of rain each day to be about .9 percent. For the second question, I found that there is a little less than a .8 percent chance of it raining 8 or more times in a month when there was a 10 percent chance each day. For my personalized histogram data, I found that it will rain an average of 6 days a month when there is a 20 percent chance each day. Lastly, I have 95 percent confidence that it will rain between 0 and 8 days in a month.

I. INTRODUCTION

This project was meant to challenge us to take everything we've learned this year, and use it to solve a practical problem, given initial conditions. In specific, it tests how well we know how to complete a problem by use of the Monte Carlo method. The Monte Carlo method is a method in which you find the probability of something happening in given circumstances by simulating the circumstances many times. The test we were given was first to determine what the probability of rain on just one day was, given that the probability of rain was $1/5$. Then, we had to find out what the probability is to rain eight days, with a lower probability to rain on each day. Lastly, we were directed to calculate the probability of an average rainfall of 10 cm, given several initial conditions for the probability of it raining, and how much each day. We then had to graph it on a histogram, average it, and find the uncertainty.

The approach that we were expected to take to solve this was the Monte Carlo approach. Basically, I had to write a code that took in the given probabilities, simulated however many months I told it to, and then tell me what the percent chance was of the proposed condition coming true.

II. QUESTION 1

First, I figured out this problem numerically. The question was asking what are the odds that it does rain the first day AND doesn't rain the second day AND doesn't rain the third day and so on. Then you calculate the "or" part. What it is the probability that it will do what I just stated, OR what is the probability that it won't rain the first day AND rain the second day AND won't rain the third day and so on. So, you must do $((.20 * .80^{29}) * 30) * 100$. This gives you a percentage of .928. Before I did any coding, I had to import numpy. I then set a variable equal to the number of days in a month, which was defined to be 30. Next, I made a function that would tell me how many times it rained in a month. Within the function, I set a counter equal to zero. I then made a loop for a day in the range 0 to 30

days. Within that loop, I generated a random number for each day. I proceeded to write a conditional that would add a number to my counter every time the random number that was generated was below .2. The function would then return the number in the counter.

That's not all though. I then made a function to tell me how many times the counter was equal to one. I set a variable equal to the number of months I wanted to simulate. Then I made a counter that would tell me how many times this the previous function returned 1. I made another loop, that ranged from 0 to however many months I chose, and when the previous function equaled 1, a number would be added to my new counter. The function then returned this new total and divided it by the number of months. This number, times 100, is equal to the percentage of times that it only rained once in a month.

III. QUESTION 2

For question 2, the condition changed from a 20 percent chance of rain to a 10 percent chance, and I was asked to calculate the probability of it raining at least 8 times in a month. This coding was the same as the last one basically. The only difference was, in the first function, instead of the random number being less than or equal to .2, it now had to be less than or equal to .1.

Then, in the second function, I changed the conditional to add a number to my new counter if the value of the first function was greater than or equal to 8. Everything else was exactly the same.

IV. QUESTION 3

A. 3A

I couldn't figure out 3A.

B. 3B

Because I couldn't figure out 3A, I based my histogram off of my first group of data. When there's a 20 percent chance of rain for each day, how many days will it rain in a month? I made a list called values, and appended it with 1000 results of my function from question 1. I had it plot the values from the list, using 30 bars, and I got rid of the lines in between. I titled it Monthly Rain.

C. 3C

To find the average amount of days per month that it rains, I created a new function. I set a counter equal to

0, and for all values of my original function that were greater than zero, that value got added to the counter. I then divided that number by the number of months.

D. 3D

I'll admit, I really didn't understand how to do this one by code either. I found the sort function on the internet, so I subtracted the heighest and lowest 25 numbers of my data, and then averaged each group. I got the bounds to be 0 and 8, and I made a string saying so.