Hypothetical Rainfall Data

H. Agnew

Siena College, Loudonville, NY

(Dated: December 2, 2016)

I did this final project of predicting the rainfall levels of hypothetical months in various scenarios in order to hopefully maintain my grade and represent weather data based upon given parameters. I did this by throwing several hours at my laptop with my coding knowledge and in the end, it ended up working the way I wanted it to, so thats pretty exciting.

I. INTRODUCTION

This project used the computer language Python in order to create a rainfall simulator for various situations and parameters. This was done mainly by using a Monte-Carlo approach, in which you use the generation of many random numbers and use those numbers in a function in order to get a reliable answer even though the individual inputs are generated at random. In our first two scenarios, we used a basic month with only one constraint and then for our third scenario we used several constraints in order to create a more realistic representation of actual rainfall values.

II. SCENARIO ONE

In my first scenario, the objective was to determine the odds that it rained one and only one day in a hypothetical month of thirty days, given the constraint that the odds that it rained in a specific day were one in five. In order to do this, I generated a random number in between zero and one, then took that number and if it was less than one fifth, assumed that it had rained that day. In order to get the data for the full month however, I created another function that went over the initial day 30 times, the length of our hypothetical month, and kept a count of how many times it had rained. Finally in order to visualize the data and draw conclusions from it, I simulated the month 1000 times and plotted the data in a histogram. In this plotting process, whenever a month had rained only one day that month was added to a counter in order to determine the percent chance that a month would rain only for a day, and this percentage was determined to be approximately nine tenths of a percent, which is the percentage you would receive if you calculated the probability out as such, $(.2)(.8^{29})(30) = .093$ as two tenths is the probability for rain, and four fifths to the twenty-ninth power is the probability for no rain and thirty is for the thirty possible combinations allowing only one day of rain.

III. SCENARIO TWO

In my second scenario, the objective was was to determine the odds that it would rain at least eight days in another hypothetical month of thirty days but this time the constraint was that the odds of rain in a day were one in ten instead of the previous one in five. In order to do this I took a similar approach, in which I generated another random number, but this time if the number was less than or equal to one tenth then it was to be assumed that it had rained that day. Next, again, I defined a function in order to go over thirty days and keep a count of the days that had rained. Then, in order to visualize the data and draw conclusions again, I plotted the 1000 simulated months in another histogram, and kept a counter of the hypothetical months in which it had rained at least eight days. Finally I used that counter in order to determine that the odds in which it would rain at least eight days in a month given the parameters to be eight tenths of a percent.

IV. SCENARIO THREE

In the final scenario, the objective was to determine the rainfall data and uncertainty values for a month in which there were several constraints in order to make the final data more realistic. Also, if it had rained in a day, there were multiple possibilities for the amount of rain, again, to make the data more realistic. In order to tackle this task, I coded the probability for each possible option in the first three days, as every day afterward only requires the data from the previous three days in order to determine if it would rain on that day. I did this by setting up a series of nested conditionals that would lead you down a plinko-esque pathway that was determined by the generation of random numbers and the probability for rain on each possible path. Once I had created a pathway for the first three days it was possible to code for every other day at once. I did this by taking each possible option from the previous function and using that to determine the probabilities for rain and looped over this to finish off the hypothetical month, this was done all while keeping a count of the quantity of days with rain. Once I had that data, I graphed it in order to get a better visualization of the data, but with that data it was possible for me to take each rainy day in the hypothetical month and use another randomly generated number in order to determine the amount of rain on that day based upon the odds for each possible amount of rain. After I had my rainfall data, I used the data in

order to plot a histogram of many simulated months, determine the average amount of rain, which I deemed to be around 7.3 centimeters, and determine the chance that it rained more than 10 centimeters in a hypothetical month, I determined this chance to be approximately 32 percent. Finally I took each individual simulated month that was used in the plotting process, sorted them by magnitude and then cut off the outliers of the final 2.5 percent of each extremity in order to get my uncertainty values, which allowed me to say that I was 95 percent sure that it would rain somewhere between zero and seventeen centimeters.