Math 1780.001 (Cherry) Computer Project #3

Due: at the start of class on Tuesday, August 11.

Project Reports must be typed! and may include computer print-outs.

Purpose. The purpose of this project is to illustrate the Weak Law of Large Numbers and the Central Limit Theorem.

Part I. Coin Flips

We will use Excel to simulate coin flips.

1. Recall that we can simulate a coin flip in Excel with =IF(RAND()<0.5,1,0). Enter this formula in cell A1.
2. In cell A4 enter the formula =AVERAGE(A1:A2). This will make sense later, but for now you should just see cell A1 repeated.
3. Select cells A1–CZ4 and choose “Fill Right” from the “Fill” tab. This should give you about 80 coin flips in row 1 and repeated in row 4.
4. In cell A10 enter the number 0.
5. In cell A11 enter the formula =A10+0.1.
6. Select cells A11–A20 and select “Fill Down.”
7. In cell B10 enter the formula =A10+0.05.
8. Select cells B11–B20 and select “Fill Down.”
9. Select cells C10–C20 and array enter the formula
   \[=\text{FREQUENCY}(a4:cz4,B10:B20)/(0.1\times\text{COUNT}(a4:cz4))\]
   Don’t forget to enter the formula while holding down the control and shift keys to enter the formula as an array. The \text{COUNT} function just counts the number of numbers in the cell range. The 0.1 comes from the bin width, so the denominator is exactly what we should divide by to convert the output of the frequency function into a probability distribution, so total area under the curve is 1. Remember to get area, we would take bin width times height for each bar and add them up.
10. Select cells C10–C20 and insert a column chart to get a histogram of the coin tosses. To fix the values on the horizontal axis, select the horizontal axis with the mouse and with the right mouse button choose “select data”. Select “Edit” from under “Horizontal axis (category) labels”, and then select cells A10-A20 with the mouse. You should see only two bars, one at 0 and one at 1. If you want, select the chart and using the “Layout” option from “Chart Tools”, add a title to this chart. You can call it “Weak Law of Large Numbers.”
11. Recall that the mean and standard deviation for a fair coin toss are both 0.5. In cell A5 enter the formula
   \[=\sqrt{\text{COUNT}(A1:A2)}\times(A4-0.5)/0.5\]
   You’ll see the point of the \text{SQRT} (\text{COUNT}()) part in a bit. Note that the 0.5’s that appear in this formula are from the mean and standard deviation. You will need to change these numbers in the later parts of this project when you work with other distributions.
12. Select cells A5–CZ5 and fill right. Row 5 now contains data for the variable \(Y_n\) on page 311 of your book.
13. Enter the number −3 in cell D10.
14. In cell D11 enter the formula =D10+0.2.
15. Select cells D11–D40 and fill down.
16. Select cell E10 and enter the formula =D10+0.1, and then select cells E11–E40 and fill down.
17. Select cells F10–F40 and array enter the formula
   \[=\text{FREQUENCY}(a5:cz5,e10:e40)/(0.2\times\text{COUNT}(a5:cz5))\]
   We’ll use this for a histogram of the \(Y_n\) variables.
18. In cell G10 enter the formula
   \[ =\text{NORMDIST}(D10,0,1,\text{FALSE}) \]
   This gives the theoretical probability distribution for the standard normal distribution.

19. Select cells G10–G40 and “Fill Down.”

20. Select cells F10–G40 and insert a column chart. This will put two histograms on the same chart. Fix up the horizontal axis by selecting the horizontal axis, using the right mouse button to choose “select data.” Select “Edit” from under “Horizontal axis (category) labels”, and then select cells D10–D20 with the mouse. You should see two sets of colored bars. One set should show two bars (one at 1 and the other at −1) and the other should be the theoretical bell curve. If you want, add the title “Central Limit Theorem” to this chart.

The idea now is to add more and more coin tosses. The first histogram we created will illustrate the Weak Law of Large Numbers and the second histogram will illustrate the Central Limit Theorem.

21. Select rows 1 and 2 and enter “Fill down.” Thus we will now average two coin tosses in each column. You should see your histograms change to have three bars.

22. Now select row 2 and choose “Insert” from the “Cells” group on the “Home” menu.

23. Now select rows 1–2 and choose “Fill Down.” We now have 3 coin tosses in each column.

24. Now select rows 2–3 and choose “Insert”. Then select rows 1–3 and select “Fill Down.” We now have five coin tosses per column.

25. We can continue to double the number of coin tosses in each column in this manner.

26. You should notice the bars in the first histogram to begin to cluster around the mean. This illustrates the law of large numbers.

27. The second histogram should show the experimental histogram start to look more and more like the theoretical bell curve. This illustrates the central limit theorem.

Print out some histograms of both types for increasing numbers of coin flips in each column to turn in with your project report.

Part II. Bernoulli Distribution. Repeat the above, but use probability \( p = 1/3 \) instead of 1/2. Don’t forget to adjust the mean and standard deviation as appropriate in the formula that compares to the standard normal distribution.

Print out histograms to include with your project.

Your project write up must explain exactly what formulas you used in your spreadsheet. You do not need to print a whole sheet of formulas, just describe the general set-up.

Part III. Dice Rolls. Repeat the above, but this time using dice rolls instead of coin flips. Remember a die roll can be simulated by \( =1+\text{INT}(6*\text{RAND}()) \). You’ll need to replace the “bin array” for the first frequency function with something more appropriate for dice. Don’t make too many bins; 10–30 seem to work well. Don’t forget to describe the bins you chose for your frequency function when you describe the formulas used for this part. Also, don’t forget that you’ll need to compute the mean and standard deviation for a die roll. Again, all calculations and spreadsheet formulas must be described in your project write-up.

Part IV. Uniform Distribution. Repeat the above, but this time using just \( =\text{RAND}() \) to simulate a uniform distribution. Describe your spreadsheet formulas in the project write-up.

Part V. Exponential Distribution. You can use \( =-\ln(1-\text{RAND}()) \) to simulate an exponential distribution with mean 1 and standard deviation 1. Repeat the above for this exponential distribution. Again, set the “bin array” for the first frequency function to something appropriate for an exponential distribution. Describe your spreadsheet formulas in the project write-up.

Extra Credit: Geometric Distribution. For extra credit, repeat the above for the geometric distribution.

A related example: The website http://wcherry.math.unt.edu/math1780/proj3 contains a spreadsheet that illustrates this project done for a four-sided die (you are not asked to do this example, but you can see the general format of how things should look).