Computer Lab Project No. 4

Exploring Probabilities

Today's lab will allow us to further explore the concept of probability through the use of *simulation.*

Info

Simulation allows you to replicate or produce data similar to the way a true physical random process would behave. We will consider two types of random processes here.

1. **Discrete uniform:** Here, you can specify a minimum and a maximum integer, and random numbers in this range will be created, so that each one occurs with the same probability. You can specify how many rows and columns of random numbers you want.

   To produce such a random data set, click "Data → Simulate → Discrete Uniform". Under "Seeding" it should say "Use dynamic seed"; this will make the results more random.

2. **Bernoulli:** Here, you can specify a probability p (meaning the probability of a "success"), and a random list of 0s and 1s is created, so that 1 occurs with probability p and 0 occurs with probability $1 - p$. Again, you can specify how many rows and columns of random data you want.

   To produce such a random data set, click "Data → Simulate → Bernoulli". As before, the seeding should be left at "use dynamic seed".

Do now

1. Simulate the result of tossing a fair coin 152 times, generating a list of 0s and 1s, with 152 entries, where 0 represents tails and 1 represents heads. Since the coin is supposed to be fair, you can use a uniform distribution with minimum 0 and maximum 1. Alternatively, you can use a Bernoulli distribution. What should the parameter p be?

   Use your simulation to estimate the probability of getting tails. Is is reasonable to expect at least 127 tails in 152 coin tosses? Why or why not? You may make a frequency distribution in order to analyze the results of the simulation.

2. Shaquille O'Neal is a professional basketball star who had a reputation for being a poor free throw shooter with a success rate of 0.528. Simulate the results of 200 free throws by creating a list with 200 entries, each either 0 (for a miss) or 1 (for a hit). Think about what distribution is appropriate here.

   Repeat the simulation of free throws five times and record the number of times that the free throw was made. Is the percentage of successful free throws from the simulation reasonably close to 0.528 in each case? You may use an appropriate frequency table for this. Study the sequences of hits and misses, how long is the longest run of misses? How long is the longest run of hits? Compare this with your classmates.

3. The probability of randomly selecting an adult who recognizes the brand name of McDonald's is 0.95. Conduct a simulation of size 10 and record the number of consumers who recognize the brand name of McDonald's. Is the proportion of those who recognize McDonald's reasonably close to the 0.95? Try another simulation this time with sample size 75. How do the results compare?

4. Lastly, let's consider a well known problem from probability theory called *The Birthday Problem*. Originally, it asks to find the probability that in a class of 35 students, at least two have the same birthday.

   Similarly to number 1 above, simulate 35 random birthdays (Hint: you can view a birthday as a number between 1 and 365, each having the same probability of occurring).

   To see whether a birthday occurred more than once, you would now have to go through this random list of 35 numbers between 1 and 365, and see whether one of them occurs at least twice. To make life easier, instead make a frequency table of the random numbers generated. When doing this, at the very bottom of the window, check "Store data in table" under "Output." Don't turn on binning. This will create two new columns; one containing the distinct "birthdays," and one containing the frequency with which each one occurred in the simulation.

   Rename the "Frequency" column "F".

   Now make a frequency table of the "F" column. Don't include that one in the table. This table will tell you which frequency occurred with which frequency. Try to understand what that table says about the random birthdays that were generated.

   Include that table in a document that you submit on Blackboard, together with an interpretation of what this frequency table means.

Finish

- Create a pdf document containing:
  1. the table listing the distinct birthdays with their frequencies,
  2. the table listing the distinct frequencies with their frequencies,
  3. an explanation (in your own words) of the meanings of these tables.
- Name this pdf file "Lab4-Firstname-Lastname.pdf" and submit it via Blackboard as your Lab 4 assignment.