**Simulation and Probability Teaching Notes**

Most students (and even teachers) find elementary probability theory challenging. Also, unlike most other high school mathematics, it is challenging for students to “show their work”. The approach in this chapter is to give students an introduction to uncertainty via simulation, rather than probability theory. The first few simulations are simple (and perhaps relatively uninteresting contexts) in which students get used to the language of simulation and become familiar with some functions in Excel necessary to do simulations. The middle of the chapter has some “juicy” questions in which different simulation techniques are used. The chapter ends by working on a great problem for a number of days - The German Tank Problem. The German Tank Problem has no “right answer”, so it is a great way to end a chapter that is based on simulation. Depending on how much time you have, there are suggested projects at the end of the chapter.

Problem 8-0

The flip “real coin” – “imaginary coin” exercise is a great way to start the chapter. It works best if it is assigned before there is any introduction to simulation/probability. When you assign the problem, be sure to stress how important it is to for them to follow the instructions. If you don’t some students will flip 5 coins at a time to save time. Tell them to print out the sheet at the end of the problem set that has a table for the 100 Hs and Ts; you could decide to print it out for them, photocopy and distribute. As students come into the room, ask them to give you their table for this question with their name on it, but not if they did it with a real or imaginary coint. The algorithm to guess whether they used a real or imaginary coint is to look for runs of 5 or more. Most students who are faking it will not put any runs of 5 or more because they think that it will not look random if they do. In fact only about 5% of the time will a random flip of a coin 100 times result in all runs being 4 or less. There is a Teacher Only Excel spreadsheet that simulates the result. Of course students will ask how you did it. The solution is left to them as one of the projects at the end the chapter.

Relative Frequency and Simulation (PowerPoint)

This PP is a good way to set up problem set 8-1.

Problem 8-1-1

Before you start this problem, you might answer the question by rolling a pair of dice. After one roll, the relative frequency is either 0 or 1, which motivates the idea that “as the number of trials approaches infinity, the relative frequency approaches the probability. We need to do a lot of trials, so let’s use Excel to do a simulation. There is an Excel solution to this problem posted for students to see. In order to help students with this first simulation, there is an avi movie file they can watch that takes them through step by step.

After students have completed this problem, show the hidden Probability of "sum equals 7" (PowerPoint). It will confirm that the relative frequency matches the probability quite closely.

Problem 8-2-2

Pose the question, “In a family with three children, what is the probability that the family will have exactly 1 boy and two girls”? Have the class discuss the question and the answer. Often two groups will form, one that thinks 3/8 is correct and one that thinks that ¼ is correct. Many students will think that because there are four events in the sample space that the probability must be ¼. I suggest you let them “argue” about the correct answer for a while. Let the simulation “settle the argument”. A simulation is posted that gives the answer. You might want to hide the solution until they have completed the homework. After they have completed 8-2-2-b, show the Family with 3 Children - 8 Equally Likely Outcomes\* (PowerPoint). It will confirm that the relative frequency distribution matches the probability distribution quite closely. Mention that the answer ¼ was not correct because the 4 outcomes were not equally likely. You can reflect back to the 36 events associated with the sum of two dice; these events *are* equally likely.

Problem 8-3-1

I would suggest that you *not* see if there is a match of two birthdays in your class before you do 8-3-1. Too often there is a match and if there is, it ruins the surprise when they find out how often there is a match. There is a posted Excel solution to the problem that I suggest you hide at first. The trick to this simulation is using the Excel function “mode”. Note: this simulation is different than the previous ones because recalculating (pressing F9) only gives you one trial. Students do 100 trials for homework, tabulating their results. In class, put all of their relative frequencies together to get one relative frequency for the entire class. There is posted *advanced* solution that allows you to do many trials of the simulation quickly.

Problem 8-3-2

The elevator problem can be done as a second problem or could even be used as a quiz.

Problem 8-4-1  
The cracked egg problem is a binomial probability question, but it is not presented to the students as such. The fact that students can answer questions about binomial probability in the fourth section of the chapter attests to the power of simulation. Get them to answer 8-4-1-d as a cell in Excel (a dynamic answer, not a static one), that include the cells for exactly 2-12 cracked eggs. There is a posted Excel solution to this problem.

Problem Set 8-5

To start, ask students how they would *jumble* the integers 1-10. Almost always they will say to use =randbetween(1,10). Tell them to try it and the will quickly see that integers can be repeated. That motivates the need for another algorithm. The movie explains how to create the simulation. One of the key ideas in this problem is that in part b it would be advantageous to report the relative frequency of 5 or more women being laid off, where as the company in part c would report the relative frequency of exactly 5 women being laid off. The spreadsheet called “Layoffs – advanced” gives the teacher a way of doing many trials quickly.

Problem Set 8-6

Start by having students read the article. As it says in the problem set, our job is to make sense of:

22% approve - Cell B2 =IF(A2<=0.22,"A","D")

1,100 people polled – There are 1,100 rows in columns A and B corresponding to the 1,100 people being polled

Margin of Error: *±* 3 – 95% of the ratings will be within 95% of 22%.

95% of the time– Run the simulation 200 times (either using student data or by running “Approval Rating Poll – advanced”) and put the relative frequency of those who approve is ascending order. Throw out the smallest 5 and largest 5 approval ratings. Then you have 95% of the ratings

Problem Set 8-7

The German tank problem can be done over a number of days. I give a suggestion of how you might do the problem in five days, though it could be done in 3 or 4.

Day 1  
I like to tell the story of “The German Tank Problem” as a lead off instead of having them read about it. There are many places on the web to read about it. I lead them through part 8-7-1-a in class. If they can’t make a guess after being given just one serial number, ask them the following. Suppose you are on a game show and there are 40 ping pong balls in a hopper numbered 1-40. The game show host draws one ball at random and you and the other contestants must guess the number of the ball. The contestant whose guess is the closest wins the car. What number would you guess? Almost everyone will say 20 right away. (Note: At first, most will think of 20 as the average of the integers 1-40, but it isn’t – 20.5 is. You can leave that detail go for now, because it is addressed in 8-7-1-c.) After that conversation, most will agree that doubling the first serial number is a good estimate, because we think the first serial number is “in the middle”. Connecting back to Chapter 1, the univariate measures of center are mean and median. This leads to the estimator 2\*mean and 2\*median. The movie for this section shows how to create the simulation that includes these two estimators. This takes you through 8-7-1-a & b.

Day 2

In 8-7-1-c, students are supposed to discover that 2\*mean-1 and 2\*median-1 are improvements to the estimators used so far. What is interesting and probably surprising about 8-7-1-d is that you need about 20-30 serial numbers before one can be relatively confident, irrespective of the total number of tanks. There are a number of ways to define “confident”. Students might be satisfied to leave the notion of confidence at the qualitative level. If they want to be more quantitative, you could suggest that one would be confident if the estimate was within 5% or 10% of the actual.

Day 3

In 8-7-1-e, students should discover that 2\*mean-1 or 2\*median-1 have a built in problem. Consider the example in which there are 100 tanks total and the captured serial numbers are captured in the following order: 100, 1, 2, 3. The estimators 2\*mean-1 or 2\*median-1 will quickly estimate below the maximum. But we know that the estimate must be  the maximum. Max + min is an estimator that takes care of this problem. Other estimators that can be easily found on the web are:

Max + Min or Max + Min – 1   
Max + Max/n or Max + Max/n – 1

Median + IQR (This one and the next one connect nicely to Chapter 1)

2\*IQR

Day 4

Finally, students are ready to “attack” the situation in which they do not know the total number of tanks. This is done in 8-7-1-e. The answer is 477.

Day 5

An assessment. This can be done individually or in groups of 2 or 3. Decide on some number of total tanks (I suggest something between 500-2,000) and do not tell the students. Chose 50 serial numbers randomly chosen from the total number of serial numbers. Print them out, one serial number per student or group and place them in the same order, face down. Students/groups start with 100 points and can request as many serial numbers as they want, but they loose one point for each serial number they request. When they feel they have enough information, they turn in their estimate (an integer), but loose 1 point for the absolute difference between their estimate and the correct answer. Normalize the scores (find a z-score for each) and then multiply by a typical standard deviation of grades and finally add a typical mean. This way you don’t need to worry if the scores are too high or low relative to typical grades for your school.

.

Projects

There are five projects and there is a Teacher Only solution to all of them.