SM3 HW 12.2 Simulation

OBJECTIVES:

- Decide if a specified model is consistent with the results from a given data-generating process like using simulation

VOCABULARY:

- **Theoretical Probability** is the ratio of the number of favorable outcomes to the total possible outcomes.
$$P(event) = \frac{number\ of\ favorable\ outcomes}{number\ of\ all\ possible\ outcomes}$$
For example, if we roll a die (six sided) 60 times, in theory the results should be 10 ones, 10 twos, 10 threes, etc.
- **Experimental Probability** is the ratio of the number of favorable outcomes that resulted from the trials to the number of trials we attempted.
$$P(event) = \frac{number\ of\ favorable\ outcomes}{total\ number\ of\ trials}$$
For example, the experimental probability of those 60 die rolls would be the actual number of each number being rolled out of the 60 rolls we made. Most of the time the experimental probability doesn't match the theoretical probability exactly.
- **Law of Large Numbers** the more trials that are conducted the closer the experimental probability will become to the theoretical probability.

Simulations are an effective tool in approximating the probability of events when it is difficult to determine a theoretical probability or to conduct an experiment to establish it's experimental probability. Simulations can use any random method of generating results as long as it fits the situation. The random method must be able to model the same number of outcomes as there are in the actual event we are trying to simulate. Random digits tables are well suited for simulations because digits can be assigned to any number of events. Flipping coins are good for events with two outcomes, selecting cards from a deck or using a calculator to create random numbers is also an effective simulation tool.

Let's imagine we want to figure out the probability that Jody, who shoots 70% for any given free throw will only make exactly 5 out of the next 10 free throws.

Steps in conducting a simulation:

1. Choose your random method for determining the simulated outcomes.
   Example: I will use the random digits generator on my calculator
2. Assign outcomes in the simulation to results of the event we are simulating.
   Example: Digits 0 - 6 will be a made free throw, digits 7, 8 and 9 will be a missed free throw. It is important to notice that the digits that represent the made and missed free throws match the likelihood of that event occurring exactly. There are 7 digits that indicate a make and 3 digits that indicate a miss which match the 70% probability for making any single free throw given in the description
3. Based on the event we are simulating go through the number of trials needed and record the "successes" and the number of trials attempted. The number of trials to be attempted should be chosen before you start. This is called the "stopping rule".
   Example: We want to find the probability of getting exactly 5 out 10 made free throws so one trial would be 10 free throws, a "success" would be when we get exactly 5 made free throws, everything else would be considered a "failure". So I would look at 10 digits at a time and count how many of them were between 0 – 6. If there are exactly 5 numbers between 0 – 6 it would count as a "success"

Digits Generated	1040593826	4823997816	2938481627	2884901977
Conclusion	8 makes = "failure"	5 makes = "success"	7 makes = "failure"	4 makes = "failure"

4. Calculate the probability from your simulation.
$$P(Jody\ makes\ exactly\ 5\ out\ 10\ free\ throws) = \frac{number\ of\ favorable\ outcomes}{total\ number\ of\ trials} = \frac{1}{4}\ or\ 0.25$$

Based on our simulation we found that Jody will make only 5 out of 10 free throws about once out of every four times.
How reliable do you think the results of this simulation are?
What can be done to improved our simulation to give us more confidence in our probability?

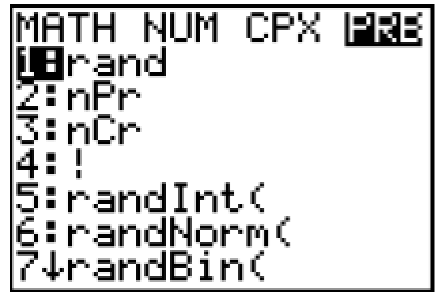
SIMULATION METHODS:

Let's look at how we could use different simulation methods to model a probability question. The school newspaper wants to conduct a poll of 20 students about the school's dress code. The journalism teacher is surprised to see that the "random sample" of students has 15 boys and 5 girls when he knows that there is a 50 – 50 split of boys and girls at the school. He decides to do a simulation of selecting 20 students at random and see if getting 15 or more boys in a random sample of 20 students is cause to question the selection process. He chooses to do 40 trials.

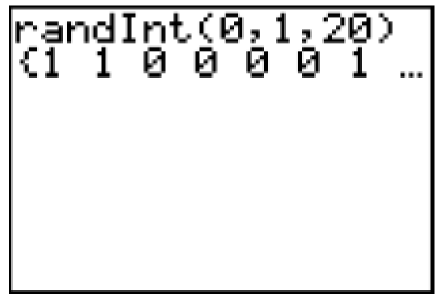
1. Flipping a coin. Since the probability of choosing a boy or a girl is 50 – 50 a coin would be a good simulation method. We would assign heads to girls and tails to boys and count up how many we got of each out of 20 flips. He would have to do 20 flips 40 times to complete his simulation.
2. Use a random number generator (Calculator). We'll assign a 0 to be a boy and a 1 to be a girl.

   On the calculator go to the MATH menu 

   . With the arrow keys go to PRB and then down to number **5:randInt(**



   Enter the values you want the calculator to generate your numbers with. The syntax is:
   randInt(lowest number, highest number, how many numbers you want generated)
   We want to use 0 and 1 and we want 20 generated each time so it would look like the example to the right.
   Hitting enter again will give a new set of 20 numbers, use the arrow keys to move right and see the other numbers in the list.



3. Random digits table. Since we have two equally likely outcomes we have to assign digits to represent those outcomes making sure that probability of each one remains 50%. We could assign 0-4 to the boys and 5-9 to the girls. We could also stay with 0 is boys and 1 is girls and ignore everything else but that would mean a lot of extra digits to deal with. The easiest would probably to make the even numbers including 0 be the boys and the odd numbers be the girls. Number tables are often separated into columns to make them easier to read, it does not have an effect on the meaning of the numbers. You can randomly choose a row to begin with but should go left to right through the whole row and then to the next one. You should not skip around the table once you've started. Here row 7 was chosen and it is broken into groups of 5 digits to make it easier to read. The first 20 numbers on row 7 has 6 evens (Boys) and 14 odds (Girls). That would be your first trial.

Random Number Table

16291	47751	28617	43266	75692	81384	25354	78664	35358	14658
93761	93658	15455	18589	64916	51584	17368	37478	53769	62767
76772	24458	49349	26977	55973	94643	77369	44195	68696	44356
64883	45331	43386	94778	35279	46898	63253	81918	63219	57955
64686	99491	32921	21687	27593	89286	56643	81317	94334	35217
67123	54977	86575	42722	91337	84614	76229	67517	23953	43454
→ 39787	57814	17496	37277	43156	21483	44215	69351	11536	51665
87251	52193	94179	65383	26512	16476	56585	85955	25919	65346
51437	78564	57291	99419	15222	64582	62473	25812	26869	41256
69143	31827	31237	55455	47444	87593	97638	57597	68126	59583
72828	24116	42381	25452	14434	15131	53789	55711	75147	96269
86675	68946	62963	58266	54867	23988	97653	34312	31265	15965
46672	78525	64155	29222	47717	93568	65534	17878	97237	85737
24575	34765	61588	335411	57237	64314	51587	28797	46111	81988
42941	71328	39677	27853	25119	65448	84123	55469	46175	44911

12.2 Exercises NAME____

PER______

Online Simulation Applets:

Coin Flips: **http://www.shodor.org/interactivate/activities/Coin/**

1. Mr.Gilchrist never studies for his faculty meeting quizzes, when Ms. Shaw reminded him that we had a quiz the next morning Gil said he wasn't worried about it. Since it was a True/False quiz with only 10 questions on it, he boasted that he was a lucky guy and should easily get at least 8 out of the 10 correct since each question had a 50% chance of guessing the right answer. Ms. Shaw thinks Gil is full of hot air and decides to see how likely it would be for Gil to get at least 8 out of 10 right if he was just guessing on each question.

    a. Let's do a simulation where Heads = A correct answer and Tails = A wrong answer. Since there are 10 questions on the quiz each trial will be made up of 10 coin flips. Either by hand or using the probability app on the TI-83 or TI-84.

	# Correct (Heads)	% Correct
Trial 1		
Trial 2		
Trial 3		
Trial 4		
Trial 5		

    b. Did you get 8 out of 10 correct (Heads) in any of your trials? How many times?

    c. What do you think are Gil's chances to get 8 out of 10 correct if he's just guessing?

    d. How would conducting more trials than just 5 affect our simulation results?

    e. Do the simulation again but let's do 30 replications of the simulation instead just 5, record your results in the table below:

Trial #	Correct	%	Trial #	Correct	%	Trial #	Correct	%
1			11			21		
2			12			22		
3			13			23		
4			14			24		
5			15			25		
6			16			26		
7			17			27		
8			18			28		
9			19			29		
10			20			30		

    f. How did the probability of getting 8 or more correct out of 10 true/false questions change with the increased number of trials?

    g. Which of the two probabilities do you think is more accurate? Explain your answer.

    h. Based on what you've discovered in this example what is an important element of conducting simulations?

2. Sarah's favorite candy, Yummies, comes in a mix of 3 flavors, purple grape, red cherry and green apple. She often browses the web site of the company that makes Yummies to learn all she can about her favorite treat. She found out that in a pack of 10 random Yummies, about 30% should be purple grape, 30% should be red cherry and 40% should be green apple. Sarah secretly wishes that every time she opens her pack of 10 Yummies, she will get at least 5 purple grape candies. Conduct a simulation using the random number table below to determine the probability that Sarah's wish will come true. Do a simulation of 20 packs of 10 candies starting with the top row.

19223	95034	05756	28713	96409	12531	42544	82853
73676	47150	99400	01927	27754	42648	82425	36290
45467	71709	77558	00095	32863	29485	82226	90056
52711	38889	93074	60227	40011	85848	48767	52573
95592	94007	69971	91481	60779	53791	17297	59335
68417	35013	15529	72765	85089	57067	50211	47487
82739	57890	20807	47511	81676	55300	94383	14893

a. Let's use digits 0 – 2 to represent purple grape, and digits 3 – 9 to represent any candy other than grape. Record the number of purple grape candies in each of the 20 packs simulated.

	# of Grape		# of Grape		# of Grape		# of Grape
Pack 1		Pack 6		Pack 11		Pack 16	
Pack 2		Pack 7		Pack 12		Pack 17	
Pack 3		Pack 8		Pack 13		Pack 18	
Pack 4		Pack 9		Pack 14		Pack 19	
Pack 5		Pack 10		Pack 15		Pack 20	

b. According to your simulation, what is the probability that Sarah's wish will come true and she will get 5 or more grape candies in her pack of Yummies?

3. LET'S MAKE A DEAL!!!! In 1963, NBC started to host a game called *Let's Make a Deal!* Contestants were given three doors to choose from. Behind one door was a prize. After selecting one door, the contestant was shown what was behind one of the doors they did not select. The contestant is then asked if they would like to stick with the door they first selected, or switch to the remaining one.

a. Which strategy do you think would result in the best chance of selecting the winning door, sticking with the door they chose first or switching doors?

b. Go to http://www.shodor.org/interactivate/activities/SimpleMontyHall/ and play the game 20 times sticking with your first choice. Then play it 20 more times switching doors. Record your wins and losses for each method in the table below:

	Sticking	Switching	Total
Win			
Lose			
Total			

c. Based on the simulation, what is $P(winning|sticking) =$

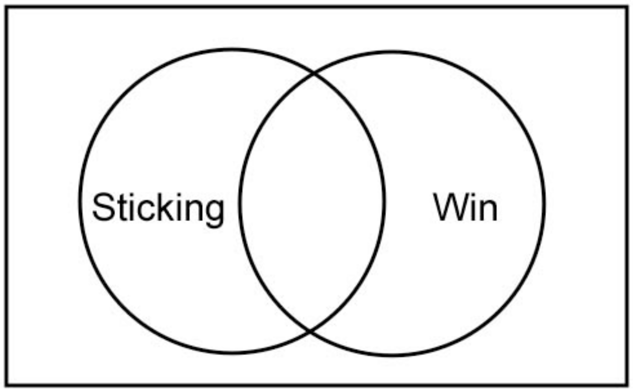
d. Based on the simulation, what is $P(winning|switching) =$

e. Does there appear to be a strategy that seems to win more often? Why do you think that method is seems to be more successful than the other?

4. PROBABILITY REVIEW
   a. Use the following table of Gil's wins and losses in his Let's Make a Deal simulation from #3 to answer the following questions

	Sticking	Switching	Total
Win	**5**	**12**	**17**
Lose	**15**	**8**	**23**
Total	**20**	**20**	**40**

b. Draw a Venn Diagram of the table data and include the probability for each area:



c. Draw a tree diagram with the probabilities of choosing a winning or losing door included.

d. $P(winning) =$

e. $P(winning \cap sticking) =$

f. $P(winning \cap switching) =$

g. $P(losing|sticking) =$

h. $P(winning\ or\ losing) =$

i. Are the events winning and sticking independent of each other? Justify your answer using probabilities.