CLEMSON MIDDLE SCHOOL MATHEMATICS PROJECT
UNIT 7: PROBABILITY

PROBLEM 1: ARE YOU LUCKY?

In your science class, you are facing a major test. Several of the questions are True/False, but, unfortunately, you must guess on four of the questions. Design and carry out a simulation to determine how many of the four you might expect to get correct.

A. Suppose you need to answer at least three of the four questions correctly. Estimate your chances for success. Write a sentence supporting your estimate.

B. Estimate your chances for guessing correctly on all four questions. Estimate your chances for guessing incorrectly on all four questions. Again, write a sentence supporting your estimate.

C. Design a simulation using coins, spinners, and/or number cubes. Carry out the simulation for thirty trials. Then carry out twenty more trials of the simulation using the random number generator on your graphing calculator.

D. Summarize the results from the fifty trials in a frequency table. Include a relative frequency and cumulative frequency column.

E. Calculate the summary statistics from the simulation. Which statistics are the most helpful? Why?

F. Create a histogram that displays your results. Comment on the histogram.

G. Based on your results, what is your probability of guessing at least 1 correct? 2? 3? All 4? How do these compare with your estimates in parts A and B?

H. Compare your results with other classmates. Then combine the class results.

I. Suggest at least three other real world examples that are mathematically equivalent to this problem.

MATERIALS

Coins, spinners, and/or number cubes

Casio *Algebra FX 2.0* Graphing Calculator

ONE SOLUTION TO PROBLEM 1: ARE YOU LUCKY?

A. Suppose you need to answer at least three of the four questions correctly. Estimate your chances for success. Write a sentence supporting your estimate.

Answers will vary, but students' responses should give the teacher insight into their sense of probability. Most should at least recognize that the probability of getting any particular question correct is 50%, but they will likely have difficulty extending their thinking into multiple questions.

B. Estimate your chances for guessing correctly on all four questions. Estimate your chances for guessing incorrectly on all four questions. Again, write a sentence supporting your estimate.

Again answers will vary, but students should recognize that their chances of getting all four correct are less than their chances for getting at least three correct. The teacher might lead a discussion here so that students realize that the probability of getting them all right is the same as the probability of getting them all wrong.

C. Design a simulation using coins, spinners, and/or number cubes. Carry out the simulation for thirty trials. Then carry out twenty more trials of the simulation using the random number generator on your graphing calculator.

Strategies will vary; many are possible. With coins, students might decide that "Heads" represents a correct guess and "Tails" represents a wrong guess. Flipping a coin four times would model one trial. More efficient students may use four coins and flip all four for each trial.

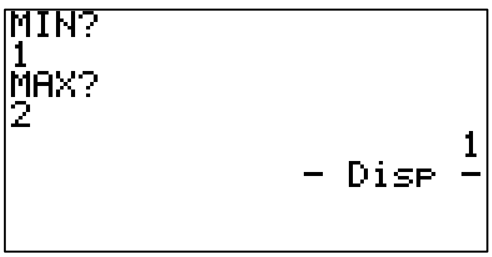
With spinners, the areas should be divided up. For example, if a spinner has eight sectors of equal area, four sectors can be used to represent a correct guess and the other four sectors, a wrong guess. If the spinner has an odd number of sectors, then one of the sectors can be considered a "non-guess," requiring a re-spin. The key is that two equal areas are used for correct and wrong guesses.

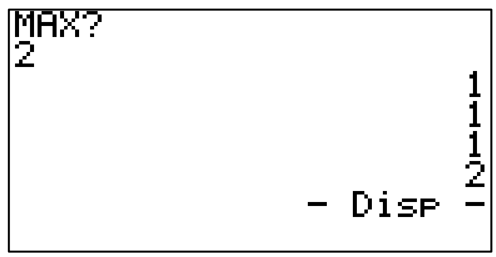
A strategy with number cubes would be similar to that used with spinners. Select three of the six faces to represent correct guesses and the other three to represent wrong guesses.

For the trials with the graphing calculator, we will have the calculator select at random either the number 1 or the number 2. We'll let 1 represent a correct guess and 2 represent a wrong guess.

If your calculator has the RANINT program on it, select PRGM from the MAIN MENU. Then,

- Highlight RANINT under the Program List and press EXE .
- The calculator displays "MIN?" Type 1 and press EXE .
- The calculator displays "MAX?" Type 2 and press EXE . See below left.
- Press EXE three more times. The screen on the right shows one trial, which represents the possible outcomes for four guesses. (Keep in mind that different people will obtain different results.) The screen shown suggests that three of the four guesses were correct.





- Simply press EXE four more times for each trial.

If your calculator does not have a program for random integers, from the MAIN MENU, choose RUN-MAT. Then,

- Press OPTN then F5 for Number. Then press 2 to obtain the Integer part of a number.
- Press the left parenthesis, F6 for more options, then F1 for Probability, then 4 for Random Number, the multiplication key, 2, right parenthesis, the addition key, 1, and EXE . See below left, recognizing that the results shown here may differ from the results shown on your screen. (NOTE: The random number generator picks a number between 0 and 1, inclusive of 0 and

exclusive of 1. The multiplication by two generates a number between 0 and 2. The integer function cuts off the decimal, leaving either 0 or 1. Adding 1 then makes the number either 1 or 2.)

x Press EXE three more times to complete a trial. For new trials, press EXE four times. Just read the bottom 4 numbers on the screen. For example, on the screen below right, read {2, 2, 1, 2}, indicating you were three out of four times. (Again, keep in mind that results will vary from person to person.)

D. Summarize the results from the fifty trials in a frequency table. Include a relative frequency and cumulative frequency column

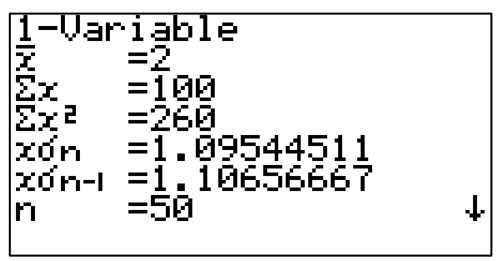
Results from one simulation are shown below. Again, answers will vary.

# CORRECT	FREQUENCY	REL FREQ	CUM FREQ
0	4	8%	8%
1	12	24%	32%
2	20	40%	72%
3	8	16%	88%
4	6	12%	100%
TOTAL	**50**	**100%**	

B. Calculate the summary statistics from the simulation. Which statistics are the most helpful? Why?

We will take advantage of the calculator to find the statistics. From the MAIN MENU, choose STAT. Then

- Clear out the data in List 1. To do so, with the cursor somewhere in the list, press [F6] for more options, [F4] for DEL-A (delete all), and [EXE] for Yes. Move the cursor into List 2 and repeat the deletion process. (If there are data in the other lists that you would rather delete, then repeat the same process.)
- Enter the number correct {0, 1, 2, 3, 4} into List 1 and the frequencies into List 2. Be sure to press [EXE] after each entry.
- If needed press [F6] for more options so that CALC is available. Press [F2] for CALC, then [4] to set up the calculations.
- Make sure 1VarXList is set to List 1 (if necessary, press [F1] with 1Var XList highlighted, then [1] and [EXE]) and 1VarFreq is set to List 2 (to do so, press [F2] , [2] , then [EXE] when 1Var Freq is highlighted). See below left. Press [ESC] to return to the primary statistics screen.
- Press [F2] for CALC and [1] to calculate one-variable statistics. See below right.



The mean tells us that, on average from the 50 trials, we had an average of precisely 2. It is very unusual for this experimental statistic to match exactly with the theoretical value. In other words, in the long run we would expect to answer two of the four questions correctly, but we should not expect to get exactly this value in a simulation. The median for our simulation is also 2. This means that half the time we had two or fewer correct and half the time we had two or more correct. The mode, the most frequent number of correct answers, is once again two. We certainly seem to have found that if we were to pick one number to represent how many we might expect to get correct, it would be 2.

However, the statistics discussed so far, while painting a consistent picture, do not paint the entire picture. We have only looked at what are called measures of central tendency. We should also look at the spread of the numbers. First we might notice that the minimum is 0 and the maximum is 4; consequently, we know that we might not get any of the questions correct, but we also might not miss any of them.

Also of interest is the Interquartile Range. The first quartile, Q1, is 1, telling us that 25% of the time we had 1 or fewer correct. The third quartile, Q3, is 3, telling us that 75% of the time we had 3 or fewer correct. This also means that we had 3 or more correct 25% of the time. If we look at the difference between these quartiles, by subtracting Q1 from Q3, we find that 50% of the time (the middle 50% of the distribution), we get 1, 2, or 3 correct.

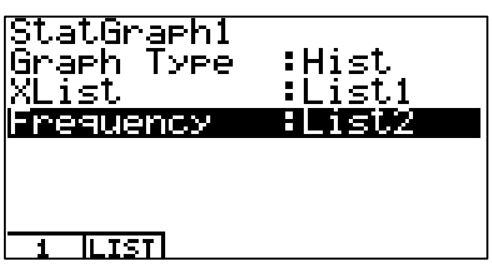
We also might note that the standard deviation, denoted by $x?n?1$ on the calculator is a little more than 1. Although this is not a literal interpretation of the standard deviation, it might help us to think that, on average, the number we will get right differs from the mean of 2 by a number greater than 1.

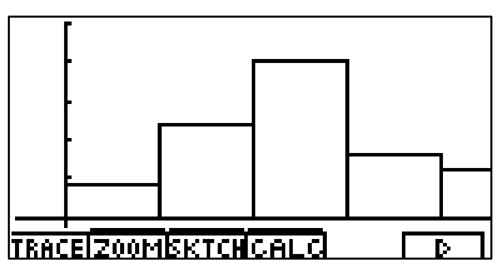
Combining the ideas of range, IQR, and standard deviation, we should begin to get a sense that even though on average we might expect to get 2 correct, we do not score 2 very consistently. The spread from 2 is significant; we should not have much confidence that on any one situation we should expect to get 2 correct. The only thing we should feel somewhat confident about is that if we guessed on four True/False questions in many, many situations, in the long run the average number we should expect to get correct is 2.

C. Create a histogram that displays your results. Comment on the histogram.

To get a histogram on the calculator,

- While looking at the statistics, press ESC to return to the primary statistics screen. Press F1 for Graph, 5 for Set.
- Use the down arrow so that Graph-Type is highlighted. Press F6 for more options and press F1 for Histogram.
- Use the down arrow and make sure the Xlist is List 1 (pressing F1 then 1 and EXE) and the Frequency is List 2 (pressing F2 then 2 and EXE). See below left.
- Press ESC to return to the primary statistics screen. Press CTRL F3 for the statistics set up menu. With Stat-Wind highlighted, press F2 for a manual window. Press ESC to return.
- Press SHIFT OPTN to set up the viewing window. Press EXE after you type in each value. A possible X-Window includes Xmin: –0.5, max: 4.5, scl: 1. Use the down arrow on the disc to bypass the dot value. A possible Y-Window includes Ymin: -5, max: 25, scl: 5. When the values are set, press ESC to return.
- Press F1 for GRPH, then 1 for S-Gph1. Type in 0 for the Start and 1 for the pitch, pressing EXE after you type each value and finally to see the graph. See below right.





The histogram shows how 2 is not only in the middle of the graph but also the most frequent value. But the graph also shows that it is certainly not unlikely to score something other than 2. However, the farther away from 2 you move, the less likely it is that you will obtain that result.

D. Based on your results, what is your probability of guessing at least 1 correct? 2? 3? All 4? How do these compare with your estimates in parts A and B?

These results are easily obtained from the relative frequency table shown earlier. We find a 92% probability of getting at least one right (only 8% of the time did we have 0 correct; alternatively, we could add the probabilities for the percents shown for 1, 2, 3, and 4 correct responses). We had a 68% probability of getting at least 2 correct (adding the probabilities for 2, 3, and 4), a 28% probability of getting at least 3 correct, and a 12% probability of getting all four correct.

Students should compare these results (or the results they obtain from their own simulation) with their original estimates.

E. Compare your results with other classmates. Then combine the class results.

Answers will vary. In the long run, the experimental probabilities should approach the theoretical probabilities. It may be informative to compare the combined results from the class with the theoretical probabilities. Using the binomial theorem, the theoretical probabilities are as follows:

?? $P(0)$? 0.0625

?? $P(1)$? 0.25

?? $P(2)$? 0.375

?? $P(3)$? 0.25

?? $P(4)$? 0.0625

I. Suggest at least three other real world examples that are mathematically equivalent to this problem.

Answers will certainly vary but some possibilities follow.

1) Suppose you have four children. How many of them would you expect to be girls (boys)? How likely is it that at least three are girls (boys)?
2) Suppose that the weather forecast calls for a 50% chance of rain for each of the next four days. How many of the four days would you expect it to rain? How likely is it that it rains on at least three of the days?
3) Suppose that on your trip to school you must pass through four traffic lights. Further suppose that the probability that any of the lights will be green when you reach the intersection is 50%. How many lights would you expect to be green when you reach them? What is the probability that at least three of the lights will be green when you reach them?

PROBLEM 2: HOW MUCH CAN YOU EAT?

A breakfast cereal includes a toy character from a popular movie in each box. An equal number of the six different characters are distributed randomly in the cereal boxes.

A. About how many boxes do you think you will have to buy to get at least one of each toy? Why?
B. It is possible to get all six characters with only six boxes. Would you expect this to happen? Why or why not?
C. It is possible not to have all six characters even after you have purchased 100 boxes. Would you expect this happen? Why or why not?
D. Design and conduct a simulation to determine the number of boxes you would expect to have to buy to get at least one of all six characters.
E. Have you ever been involved with a promotion such as this? Do you think the company is wise to do this? Why or why not?

MATERIALS

Casio *Algebra FX 2.0* Graphing Calculator

REFERENCE: "Dealing with Data and Chance," Curriculum and Evaluation Standards for School Mathematics Addenda Series, Grades 5-8, NCTM, 1997, p.15.

ONE SOLUTION TO PROBLEM 2: HOW MUCH CAN YOU EAT?

A. About how many boxes do you think you will have to buy? Why?

Answers will vary, but students should expect to purchase more than six boxes, realizing that it is highly unlikely that the first six boxes they pick will have six different characters. They should also realize that it shouldn't take hundreds of boxes of cereal until they have found all six.

B. It is possible to get all six characters with only six boxes. Would you expect this to happen? Why or why not?

As mentioned, students should recognize that it is highly unlikely for this to occur. Even if the first five boxes result in five different toys, the probability that the sixth box contains the sixth toy is only one in six. Of course, the probability is that the first five boxes will not contain five different toys.

C. It is possible to not have all six characters even after you have purchased 100 boxes. Would you expect this happen? Why or why not?

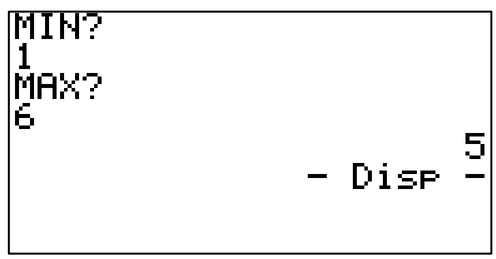
This may be more difficult for students. Nevertheless, they should be able to think through the issue. If they have not seen all six characters in 100 purchases, then, on average, they have seen *each* of the other five toys 20 times. They should realize that this does not seem very plausible either. Doesn't it seem more likely that we would obtain at least one of the missing toys before we have 20 of all of the others?

D. Design and conduct a simulation to determine the number of boxes you would expect to have to buy to get at least one of all six characters.

Let's let the numbers 1, 2, 3, 4, 5, and 6 represent the six toys. We need to have all of the numbers if we are to have all of the toys. One way to conduct the simulation is to roll a number cube and determine how many rolls it takes to obtain all six numbers.

We can also use the graphing calculator to conduct the simulation. The random number capabilities of the calculator provide us with an excellent tool. All we need do is have the calculator choose a random number between 1 and 6, inclusive, and determine how many numbers it takes until we have seen all six of the numbers. To do so, from the MAIN MENU choose RUN-MAT. Then,

- Press OPTN , then F5 for NUM (number). Press 2 for the integer component of the number.
- Press the left parenthesis, OPTN , and F6 for more options, then F1 for Probability. Press 4 for Ran#. Now type in the multiplication sign, 6, the right parenthesis, the plus sign, and 1. Finally, press EXE . See below left. This will produce integers between 1 and 6 inclusive, therefore the result you see on your screen may differ from the one shown.
- Alternatively, from the MAIN MENU choose Program. Run the RANINT program, inputting 1 as the minimum and 6 as the maximum. Press EXE to obtain a random integer. See below right, recognizing again that your results may be different.



In either case, continue pressing EXE to generate more random integers between 1 and 6. When using the random integer generator that is built into the calculator, the calculator screen will display seven numbers at a time. It can be helpful to push EXE seven times, record the results, press it seven more times, record the results, and so on. The RANINT program displays six numbers at a time. When using RANINT, press EXE six times, record results, press it six more times, record results, and so on. To escape the RANINT program, press the AC/on key, then the ESC key.

One issue students must deal with is keeping track of their results. Some sort of recording scheme is needed. Keep in mind that to conduct a single trial of the experiment, you must count how many boxes of cereal (how many times you must press EXE) until you have obtained all six toys. Of course, just doing this simulation one time will not be overly informative; we should conduct the simulation as many times as possible.

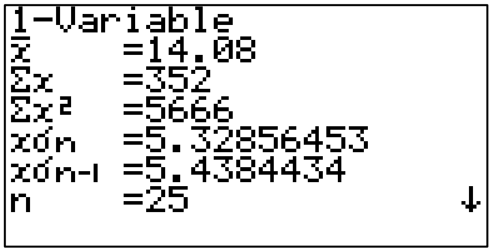
Although 25 trials is not sufficient to have much confidence in making a prediction, the results from 25 trials are shown in the table below.

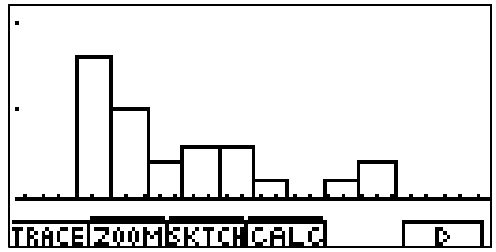
# OF BOXES	FREQUENCY	CUM FREQ	PERCENT	CUM %
8	3	3	12%	12%
9	2	5	8%	20%
10	3	8	12%	32%
11	2	10	8%	40%
12	3	13	12%	52%
13	1	14	4%	56%
14	1	15	4%	60%
15	2	17	8%	68%
16	1	18	4%	72%
17	1	19	4%	76%
18	2	21	8%	84%
19	0	21	0%	84%
20	1	22	4%	88%
21	0	22	0%	88%
22	0	22	0%	88%
23	0	22	0%	88%
24	1	23	4%	92%
25	1	24	4%	96%
26	0	24	0%	96%
27	1	25	4%	100%
TOTAL	**25**		**100%**	

Note that the minimum number of boxes purchased, according to our simulation, was 8, and that the maximum was 27. Summary statistics from the simulation and a histogram of the results may be helpful in understanding the results.

From the MAIN MENU, choose STAT.

- Delete data in Lists 1 and List 2 (and others as well, if you prefer not to see the values from unused lists). To do so, press [F6] for more options, and with the cursor someplace in the list, press [F4] then [EXE] to delete the values in the list .
- Type in the numbers from the # of Boxes column into List 1 and the Frequency into List 2. Remember to press [EXE] after each entry.(The rows with 0 frequency may be omitted.)
- To obtain the summary statistics, press [F6] if necessary to return to the first part of the primary statistics screen options, [F2] for calculations, and [4] to set the calculations.
- Make sure the 1Var XList is set to List 1 with the 1Var Freq set to List 2. When finished, press [EXE] .
- Press [F2] for calculations and [1] for one-variable statistics. See below left.
- To view a histogram of the data, press [ESC] to return to the primary STAT screen.
- You can either set the window manually, or if you are only interested in trends, you can have the window set automatically. After checking the set up, ([CTRL] [F3]) press [ESC] to return to the primary STAT screen.
- Press [F1] for graph, then [5] to set up the graph. Make sure the GraphType for StatGraph1 is a histogram with Xlist List 1 and Frequency List 2. Press [ESC] and [F1] then [1] to complete the graph. Your results may differ from the one below, depending on the start and pitch you set. The graph below is generated using *start:8* and *pitch:2.12*.





Again, only 25 trials in a simulation is too few to have much confidence in the results, but we can nevertheless begin making conjectures. The mean number of boxes of cereal needed to ensure that all six toys were received was just over 14, with a median of 12. The range was from 8 to 27, but the histogram suggests that the most likely number of boxes needed to achieve the goal is closer to 8 than to 27.

After students have conducted several trials on their own, the results from all of the groups should be combined and analyzed. Compare the overall results with those shown here.

E. Have you ever been involved with a promotion such as this? Do you think the company is wise to do this? Why or why not?

Answers will vary, but students might argue that the motivation to obtain all six characters might cause people to purchase more boxes of the cereal than they otherwise would. Having only one character might cause people to purchase one box of cereal when they might normally purchase a different brand, but having multiple characters might prove even more successful.

PROBLEM 3: MAKE IT FAIR

Consider the following games. Before playing them, determine who you think has the better chance of winning. Then, use the random number capabilities of your graphing calculator to play them several times with a partner.

A. Simulate the rolling of two number cubes by choosing two integers from 1 to 6 at random. You get a point if the sum is even; your partner gets a point if the sum is odd. The first one to reach 10 points wins.

B. The game is the same as before, but this time the product of the two numbers, not the sum, determines the winner.

C. You and your partner each begin with 10 points. If the sum of the two number cubes is 7, your partner must give you 3 of her/his points. If the sum is anything except 7, you must give your partner 1 of your points. The person with the higher point total after 10 turns is the winner. Anyone who runs out of points before the 10 turns have occurred loses automatically.

1) Compare your results with others who have played the games. Do you believe your determinations about who had the better chance to win were good ones, or should you change them?
2) Use the results to determine if each of the games is fair, that is, if neither person has an advantage.
3) Try to analyze the game. If it is unfair, change the rules to make it fair.

REFERENCE: "Dealing with Data and Chance," Curriculum and Evaluation Standards for School Mathematics Addenda Series, Grades 5-8, NCTM, 1997, pp.12-13

PROBLEM 4: MULTIPLE CHOICE

You are taking a test, but are totally lost on 10 multiple choice questions. Each question has four choices, and you are forced to guess blindly. Design and carry out a simulation that will determine how many of the 10 questions you might expect to answer correctly. Write a paragraph discussing your results.

TEXT SECTION CORRESPONDENCES

The materials in this module are compatible with the following sections in the listed texts.

TEXT	SECTION
AW – Foundations of Algebra and Geometry (1998)	7.1-2
Glencoe – Mathematics Applications and Connections C1 (1995)	2.6; 10.5; 14.1-3; 14.5B
Glencoe – Mathematics Applications and Connections C2 (1995)	3.7; 4.8; 11.8; 13.1; 13.3-4; 13.6
Houghton Mifflin – The Mathematics Experience I (1992)	5.11; 6.1; 10.2; 12.7-10; 12.12; 15.7
Houghton Mifflin – The Mathematics Experience II (1992)	1.10; 3.6; 3.8; 4.5; 9.3-4; 9.6-7; 10.6; 10.12; 10.14-15; 14.4
McDougal Littell – Gateways to Algebra and Geometry (1994)	1.4
Prentice Hall – Middle Grades Mathematics C1 (1995)	10.2-4; 10.6; 10.8
Prentice Hall – Middle Grades Mathematics C2 (1995)	1.1; 10.1-2; 10.4-6; 10.8
Prentice Hall – Middle Grades Mathematics C3 (1995)	1.3; 9.4-5; 9.7-8
SFAW – Middle School Math C1, V1 (1999)	
SFAW – Middle School Math C1, V2 (1999)	10.10; 12.1-3; 12.5-6
SFAW – Middle School Math C2, V1 (1999)	
SFAW – Middle School Math C2, V2 (1999)	8.1-2; 12.4-7
SFAW – Middle School Math C3, V1 (1999)	6.1
SFAW – Middle School Math C3, V2 (1999)	12.4-5; 12.7
SFAW: UCSMP – Transition Mathematics, Part 1 (1998)	1.6; 2.6; 4.8; 5.6
SFAW: UCSMP – Transition Mathematics, Part 2 (1998)	7.5; 13.4