LESSON TITLE: Dice and Distributions

OVERVIEW: This activity uses dice and spreadsheets to investigate empirical data (through dice rolling and data collection, as well as random simulation) as a means to introduce distributions from probabilistic concepts. Through various scenarios (single die, resulting sum of two dice, with an application to games) students strengthen their understanding of the relationship between empirical randomness and theoretical randomness while also encountering histograms/distributions. In particular, students will be exposed to the Law of Large Numbers and expected value through various exercises.

PREREQUISITE IDEAS AND SKILLS This project is intended for students who have been introduced to basic ideas in logic and probability. For instance:

- Students will need to know what probability is (and how to compute it in simple situations).

MATERIALS NEEDED TO CARRY OUT THE LESSON

- Activity worksheet
- A pair of dice (or the Excel sheet titled 'single toss simulator.')
- Excel Spreadsheet Template titled 'many toss simulator.'

CONCEPTS TO BE LEARNED/APPLIED

- Students will use learn how to use Excel's functions to create simulation of tossing dice using RANDBETWEEN, and to create histograms from the following results.
- Students will observe that the experimental average tends to gravitate toward the theoretical average, thereby introducing the Law of Large Numbers (as the number of trials of a random experiment increases, the average of the results obtained from the trials is likely to get closer to the expected value).
- Students will learn how to compute expected value, and observe (with figures) how the expected value relates to repeated experimentation of tossing dice via the Law of Large numbers.

INSTRUCTIONAL PLAN

Start the lesson with the motivating idea. This could be the following: 'Suppose you are a gambler. You want to play a game that involves random events (such as tossing dice) and there is money involved corresponding to the different outcomes of the game. You would like to make an informed guess as to whether you will make money or lose money by continually playing the game.' This will lay the ground work for students to actively engage with the idea of expectation, through empirical means.

Part 0. The instructor may use the following definition and example before allowing the students to embark on this activity.

Definition: The <u>Expected Value</u> is the average of all possible outcomes, weighted by the probabilities of those outcomes.

Example Suppose you are playing a game in which you roll a single, fair, 4 sided die. If you roll an odd number, you lose \$.10. If you roll a 2, then you make \$.60. If you roll a 4, you make \$.20.

1. What is the probability of rolling a 2? a 3?

   (Students should have been exposed to questions like this beforehand) The answer is 1/4, for both outcomes.

2. What is the expected value of tossing this die?

   The answer is:

$$\text{expected value} = 1\left(\frac{1}{4}\right) + 2\left(\frac{1}{4}\right) + 3\left(\frac{1}{4}\right) + 4\left(\frac{1}{4}\right) = 2.5$$

   We get 2.5 since it is the average of all the outcomes of rolling this die.

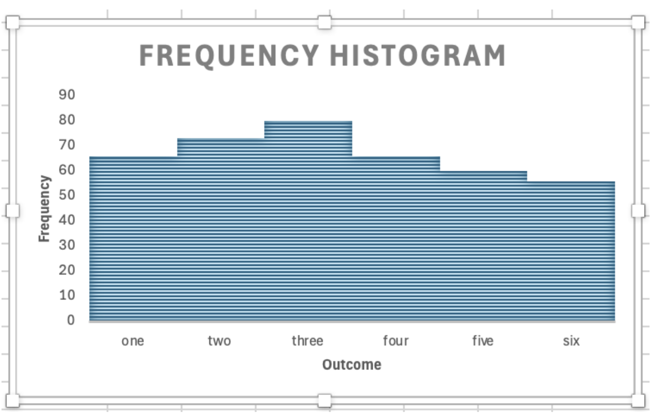
3. What is the expected payout (that is, the amount of money gained or lost) from playing one round of this game? The answer is

$$\text{expected value} = -\$.10\left(\frac{1+1}{4}\right) + \$.60\left(\frac{1}{4}\right) + \$.20\left(\frac{1}{4}\right) = \$.15$$

   which means a player is expected to make money by playing this game.

Part I. The first part of this activity explores tossing a single die. As a warm up, students will address basic probability questions such as 'What is the probability of rolling a 2?' (1/6) and 'What is the expected value of rolling a fair six sided die?' (3.5). Next, students will learn how to simulate die rolling with Microsoft Excel via the function RANDBETWEEN. Armed with the results of many tosses of the die (400, say), students will then be asked to find the average of all these results (this can be done by selecting the column with the simulated values, and reading off the Average at the bottom of the Excel sheet.) The result should be 'close' to 3.5.

The provided Excel template will create a frequency table given the simulated data. Students will be in charge of creating a histogram based on the frequency table data. (This can be done by highlighting the values in the frequency table, navigating to the 'Insert' tab, finding the drop down arrow next to 'Recommended Charts.' The formatting of the chart may be done to the instructor's preferences (or not at all). The result may look something like this:



Students will then make observations about the experimental distribution and the preparatory questions as they begin to link the mathematical probability with repeated experiments. In particular, they'll be asked to re-run the simulation (by typing SHIFT+F9) and note that both the distribution and the average will look similar with slight variations.

Finally, students will be asked to find the running average of the experiments (this can be done by using the AVERAGE function to compute the average up to the j^{th} roll, for every roll in the simulation), to be plotted and compared to the theoretical expected value of tossing a single die (3.5). This plot in comparing the cumulative average with the exact expected value will be included in the Excel template. With this figure, will observe that the experimental average tends to gravitate toward the theoretical average, thereby demonstrating the Law of Large Numbers.

Part II The next part of this project will explore the tossing of a pair of dice. Many of the components from the first part will be replicated here, with the intention of extending the ideas now to two dice. Before utilizing spreadsheets, students will be asked the following questions: 'What is the probability of rolling a pair of dice and getting a result totalling 4? totalling 5?' (3/36 and 4/36 respectively), 'Based on this observation, can you conclude that each outcome is equally likely? Why or why not?' , 'Can you hypothesize what the expected value of tossing a pair of dice is? Justify your hypothesis.'

At this point, if students are stumped, they may move on to the next question to help elucidate the reasoning. Indeed, students will then be asked to make a table displaying all possible outcomes of tossing two dice (this can be done by hand or in Excel). The students can then be asked 'Based on this table, what would you say the average outcome of rolling a pair of dice is?' and 'How does this quantity relate to the expected value of tossing a single die?'.

Once students have come to grips with some questions on tossing two dice, they can move on to the spreadsheet portion of the activity. Here, they will simulate tossing a pair of dice many times (tossing a pair of dice may be simulated by typing '=RANDBETWEEN(1,6) + RANDBETWEEN(1,6)' (without quotation marks) in the formula of a cell). Similar to before, they will then create a frequency histogram as they did in the first part of the activity. They will be asked to make comparisons of their table of possible outcomes and their experimental histograms, and will see that the distribution of outcomes is indeed not uniform. Lastly, students will once again use the Law of Large Numbers to verify that the experimental average is close to 7 after many experimental dice tosses. Once students come terms with tossing two dice, they will be asked 'Based on what you've learned about tossing 2 dice, can you write a formula for the expected value of tossing n fair six sided dice? Explain your reasoning.' in order to further generalize these ideas.

Part III In the final part of this activity, students will make sense of a simple game played with two dice. The game's rules are as follows: you roll a fair pair of 6 sided dice,

i) if the total is 2, you win \$2

ii) if the total is 7, you win \$.40

iii) if the total is 6 or 8, you must pay \$.70

iv) if the total is 12, then you must pay \$1

v) if the total is 3,4,5,9,10, or 11, you neither win nor lose money.

Before engaging with the game, students will be asked whether they think they'll be more likely to win money or to lose money by playing this game based on the rules. After the students have made their predictions, they will then be asked to play this game for at least 10 rounds (they may use real dice, or use the Excel template titled 'single toss simulator' to simulate rolling a pair of dice), and to record their results and their payout (as a loss or gain).

Once the students have participated in the playing of this game (with no real money at stake), they'll be asked to compute the expected payout of playing one round of this game. The calculation should look like:

$$\text{expected value} = \$2\left(\frac{1}{36}\right) + \$.40\left(\frac{6}{36}\right) - \$.70\left(\frac{5+5}{36}\right) - \$1\left(\frac{1}{36}\right) = -\$.10$$

They will then calculate the expected payout after playing 10 rounds ($10(-\$.10) = -\$1$) and compare their theoretical calculation to their actual outcome from playing the game. Students will finally posit what will happen after playing n rounds of the game.

MIP COMPONENTS OF INQUIRY

This section outlines how the activity will meet the Mathematical Inquiry Project (MIP) criteria for active learning, meaningful applications, and academic success skills.

Active Learning: This activity will engage students in active learning as they learn important concepts of mathematical expectations via simulation in Excel spreadsheets. Students will learn the connections between mathematical probability and repeated experimentation. For example, students will actively discover the link between the expected value of the roll of a single die (3.5) and how this relates to the average outcome of multiple rolls. Through rolling dice many times, calculating the average value, and considering how it gets closer to the expected value (e.g., 3.5 for a single die or 7 for a pair of dice), students discover the Law of Large Numbers, which states that as the number of trials of a random experiment increases, the average of the results obtained from the trials gets closer to the expected value. Through structured activities and guided questions, students will select relevant Excel functions (such as RANDBETWEEN and AVERAGE), perform simulations and calculations, and evaluate the resulting data by creating frequency tables and histograms. For examples, in the crux of the activity (part III):

1. Part III Q1: Students select an action (computing probabilities), perform the calculation, and evaluate the two numbers in order to determine whether one is more likely to win or lose the game.

2. Part III Q2/3: Students are tasked to perform an action (play 10 rounds of the game described in Part III), and then to evaluate the overall outcome of performing said action.

3. Part III Q4/5: Students select the proper form of expectation, perform the calculation, and evaluate the result as compared to the outcome of Part III Q3.

4. Part III Q6: Students must select a formula, which generalizes the result of the action of Part III Q5.

This hands-on approach will help students understand theoretical and empirical distributions and reinforce their learning through critical analysis and reflection.

Meaningful Applications: This activity will place emphasis on meaningful applications in both experimental probability, as well as the use of spreadsheets. Students will learn about theoretical probability in an empirical light. They will learn how to view real world problems in a probabilistic way (as in the gaming problem in part 3), and to interpret the simulated results in a meaningful manner (like the expected payout of a game, and whether or not they should play it to make money). In short, they will be able to use simulations in Excel to make informed decisions about a real life situation. Assessment questions throughout the activity allow students to generalize across mathematical contexts by considering alternative situations such as tossing multiple dice and playing games. Questions prompt students to make near generalizations by asking how expectation changes if another die is added, as well as far generalizations by asking them to come up with more general formulas. Further, students will learn to generalize the concept of mathematical expectation. Indeed, they will first encounter the expectation in the context of the outcome of rolling dice, and further engage with this notion by applying it to the monetary outcome of a game.

Academic Success Skills: This activity will allow students to realize the importance of careful reasoning through logical probabilistic simulations of games to model real-life situations that are non-deterministic. Further, students will learn spreadsheet tools that can be applied to many other situations (the use of functions RANDBETWEEN and AVERAGE) in order to interpret quantifiable data. This ARC will support students in developing their identities as learners by engaging them in active experimentation and reflection, fostering confidence in their ability to understand and apply mathematical concepts. Also, by tackling complex problems and interpreting data, students will enhance their problem-solving and critical-thinking skills, preparing them for future academic and real-world challenges.