Lesson 8: Theoretical Probabilities and Estimated (Experimental) Probabilities

Have you ever watched the beginning of a professional football game? After the traditional handshakes, a coin is tossed to determine which team gets to kick off first. The toss of a fair coin is often used to make decisions between two groups.

Example 1: Why a Coin?

Coins were discussed in previous lessons of this module. What is special about a coin? In most cases, a coin has two different sides: a head side (heads) and a tail side (tails). The sample space for tossing a coin is {heads, tails}. If each outcome has an equal chance of occurring when the coin is tossed, then the probability of getting heads is $\frac{1}{2}$, or 0.5. The probability of getting tails is also 0.5. Note that the sum of these probabilities is 1.

The probabilities formed using the sample space and what we know about coins are called the *theoretical* probabilities. Using observed relative frequencies is another method to estimate the probabilities of heads or tails. A relative frequency is the proportion derived from the number of the observed outcomes of an event divided by the total number of outcomes. Recall from earlier lessons that a relative frequency can be expressed as a fraction, a decimal, or a percent. Is the estimate of a probability from this method close to the theoretical probability? The following example investigates how relative frequencies can be used to estimate probabilities.

Beth tosses a coin 10 times and records her results. Here are the results from the 10 tosses:

Toss	1	2	3	4	5	6	7	8	9	10
Result	H	H	T	H	H	H	T	T	T	H

The total number of heads divided by the total number of tosses is the relative frequency of heads. It is the proportion of the time that heads occurred on these tosses. The total number of tails divided by the total number of tosses is the relative frequency of tails

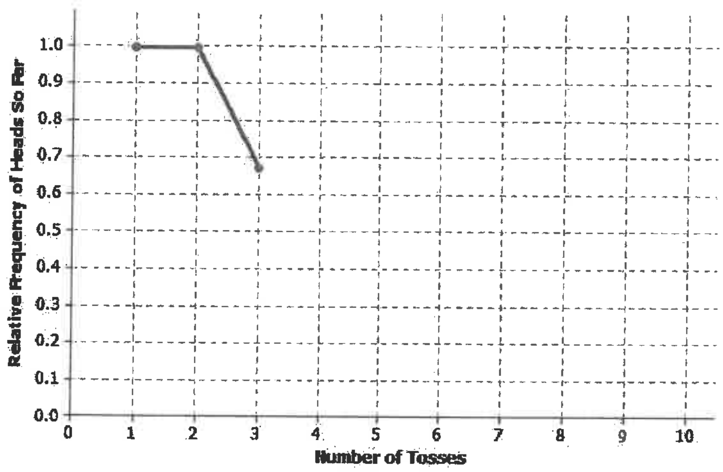
a. Beth started to complete the following table as a way to investigate the relative frequencies. For each outcome, the total number of tosses increased. The total number of heads or tails observed so far depends on the outcome of the current toss. Complete this table for the 10 tosses recorded in the previous table.

Toss	Outcome	Total Number of Heads So Far	Relative Frequency of Heads So Far (to the nearest hundredth)	Total Number of Tails So Far	Relative Frequency of Tails So Far (to the nearest hundredth)
1	H	1	$\frac{1}{1} = 1$	0	$\frac{0}{1} = 0$
2	H	2	$\frac{2}{2} = 1$	0	$\frac{0}{2} = 0$
3	T	2	$\frac{2}{3} \approx 0.67$	1	$\frac{1}{3} \approx 0.33$
4					

5					
6					
7					
8					
9					
10					

b. What is the sum of the relative frequency of heads and the relative frequency of tails for each row of the table?

c. Beth's results can also be displayed using a graph. Use the values of the relative frequency of heads so far from the table in part (a) to complete the graph below.



d. Beth continued tossing the coin and recording the results for a total of 40 tosses. Here are the results of the next 30 tosses:

Toss	11	12	13	14	15	16	17	18	19	20
Result	T	H	T	H	T	H	H	T	H	T

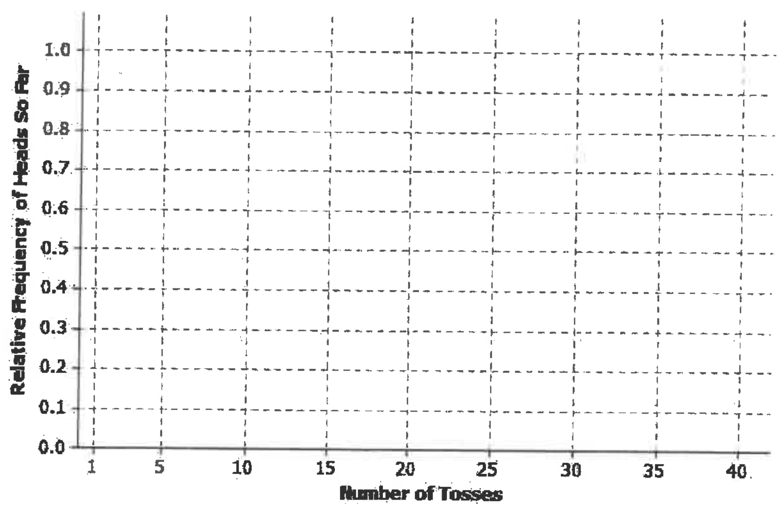
Toss	21	22	23	24	25	26	27	28	29	30
Result	H	T	T	H	T	T	T	T	H	T

Toss	31	32	33	34	35	36	37	38	39	40
Result	H	T	H	T	H	T	H	H	T	T

As the number of tosses increases, the relative frequency of heads changes. Complete the following table for the 40 coin tosses:

Number of Tosses	Total Number of Heads So Far	Relative Frequency of Heads So Far (to the nearest hundredth)
1		
5		
10		
15		
20		
25		
30		
35		
40		

e. Use the relative frequency of heads so far from the table in part (d) to complete the graph below for the total number of tosses of 1, 5, 10, 15, 20, 25, 30, 35, and 40.



f. What do you notice about the changes in the relative frequency of the number of heads so far as the number of tosses increases?

g. If you tossed the coin 100 times, what do you think the relative frequency of heads would be? Explain your answer.

h. Based on the graph and the relative frequencies, what would you estimate the probability of getting heads to be? Explain your answer.

i. How close is your estimate in part (h) to the theoretical probability of 0.5? Would the estimate of this probability have been as good if Beth had only tossed the coin a few times instead of 40?

The value you gave in part (h) is an estimate of the theoretical probability and is called an *experimental* or *estimated* probability.

Lesson 9: Comparing Estimated Probabilities to Probabilities Predicted by a Model

Exploratory Challenge: Game Show—Picking Blue!

Getting Started

Assume that the producers of the show do not want to give away a lot of their blue tokens. As a result, if one bag has the same number of red and blue chips, do you think the other bag would have more or fewer blue chips than red chips? Explain your answer.

Planning the Research

Your teacher will provide you with two bags labeled A and B. You have 20 minutes to experiment with pulling chips one at a time from the bags. After you examine a chip, you must put it back in the bag. Remember, no peeking in the bags, as that will disqualify you from the game. You can pick chips from just one bag, or you can pick chips from one bag and then the other bag.

Use the results from 20 minutes of research to determine which bag you will choose for the game.

Provide a description outlining how you will carry out your research.

Carrying Out the Research

Share your plan with your teacher. Your teacher will verify whether your plan is within the rules of the quiz show. Approving your plan does not mean, however, that your teacher is indicating that your research method offers the most accurate way to determine which bag to select. If your teacher approves your research, carry out your plan as outlined. Record the results from your research, as directed by your teacher.

Playing the Game

After the research has been conducted, the competition begins. First, your teacher will shake up Bag A. A chip is selected. If the chip is blue, all students who selected Bag A win an imaginary blue token. The chip is put back in the bag, and the process continues. When a red chip is picked from Bag A, students selecting Bag A have completed the competition. Your teacher will now shake up Bag B. A chip is selected. If it is blue, all students who selected Bag B win an imaginary blue token. The process continues until a red chip is picked. At that point, the game is over.

How many blue tokens did you win? Examining Your Results

At the end of the game, your teacher will open the bags and reveal how many blue and red chips were in each bag. Answer the questions that follow. After you have answered these questions, discuss them with your class.

1. Before you played the game, what were you trying to learn about the bags from your research?

2. What did you expect to happen when you pulled chips from the bag with the same number of blue and red chips? Did the bag that you thought had the same number of blue and red chips yield the results you expected?

3. How confident were you in predicting which bag had the same number of blue and red chips? Explain.

4. What bag did you select to use in the competition, and why?

5. If you were the show's producers, how would you make up the second bag? (Remember, one bag has the same number of red and blue chips.)

6. If you picked a chip from Bag B 100 times and found that you picked each color exactly 50 times, would you know for sure that Bag B was the one with equal numbers of each color?

Lesson 10: Conducting a Simulation to Estimate the Probability of an Event

Example 1: Families

How likely is it that a family with three children has all boys or all girls?

Let's assume that a child is equally likely to be a boy or a girl. Instead of observing the result of actual births, a toss of a fair coin could be used to simulate a birth. If the toss results in heads (H), then we could say a boy was born; if the toss results in tails (T), then we could say a girl was born. If the coin is fair (i.e., heads and tails are equally likely), then getting a boy or a girl is equally likely.

Exercises 1–2

Suppose that a family has three children. To simulate the genders of the three children, the coin or number cube or a card would need to be used three times, once for each child. For example, three tosses of the coin resulted in HHT, representing a family with two boys and one girl. Note that HTH and THH also represent two boys and one girl.

1. Suppose that when a prime number (P) is rolled on the number cube, it simulates a boy birth, and a non-prime (N) simulates a girl birth. Using such a number cube, list the outcomes that would simulate a boy birth and those that simulate a girl birth. Are the boy and girl birth outcomes equally likely?

2. Suppose that one card is drawn from a regular deck of cards. A red card (R) simulates a boy birth, and a black card (B) simulates a girl birth. Describe how a family of three children could be simulated.

Example 2

Simulation provides an estimate for the probability that a family of three children would have three boys or three girls by performing three tosses of a fair coin many times. Each sequence of three tosses is called a *trial*. If a trial results in either HHH or TTT, then the trial represents all boys or all girls, which is the event that we are interested in. These trials would be called a *success*. If a trial results in any other order of H's and T's, then it is called a *failure*.

The estimate for the probability that a family has either three boys or three girls based on the simulation is the number of successes divided by the number of trials. Suppose 100 trials are performed, and that in those 100 trials, 28 resulted in either HHH or TTT. Then, the estimated probability that a family of three children has either three boys or three girls would be $\frac{28}{100}$, or 0.28.

Exercises 3–5

3. Find an estimate of the probability that a family with three children will have exactly one girl using the following outcomes of 50 trials of tossing a fair coin three times per trial. Use H to represent a boy birth and T to represent a girl birth.

HHT	HTH	HHH	TTH	THT	THT	HTT	HHH	TTH	HHH
HHT	TTT	HHT	TTH	HHH	HTH	THH	TTT	THT	THT
THT	HHH	THH	HTT	HTH	TTT	HTT	HHH	TTH	THT
THH	HHT	TTT	TTH	HTT	THH	HTT	HTH	TTT	HHH
HTH	HTH	THT	TTH	TTT	HHT	HHT	THT	TTT	HTT

4. Perform a simulation of 50 trials by rolling a fair number cube in order to find an estimate of the probability that a family with three children will have exactly one girl.
   a. Specify what outcomes of one roll of a fair number cube will represent a boy and what outcomes will represent a girl.
   b. Simulate 50 trials, keeping in mind that one trial requires three rolls of the number cube. List the results of your 50 trials.
   c. Calculate the estimated probability.

5. Calculate the theoretical probability that a family with three children will have exactly one girl.
   a. List the possible outcomes for a family with three children. For example, one possible outcome is BBB (all three children are boys).
   b. Assume that having a boy and having a girl are equally likely. Calculate the theoretical probability that a family with three children will have exactly one girl.
   c. Compare it to the estimated probabilities found in parts (a) and (b).

Example 3: Basketball Player

Suppose that, on average, a basketball player makes about three out of every four foul shots. In other words, she has a 75% chance of making each foul shot she takes. Since a coin toss produces equally likely outcomes, it could not be used in a simulation for this problem.

Instead, a number cube could be used by specifying that the numbers 1, 2, or 3 represent a hit, the number 4 represents a miss, and the numbers 5 and 6 would be ignored. Based on the following 50 trials of rolling a fair number cube, find an estimate of the probability that she makes five or six of the six foul shots she takes.

441323	342124	442123	422313	441243
124144	333434	243122	232323	224341
121411	321341	111422	114232	414411
344221	222442	343123	122111	322131
131224	213344	321241	311214	241131
143143	243224	323443	324243	214322
214411	423221	311423	142141	411312
343214	123131	242124	141132	343122
121142	321442	121423	443431	214433
331113	311313	211411	433434	323314

Lesson 11: Conducting a Simulation to Estimate the Probability of an Event

Example 1: Simulation

If a basketball player typically makes five out of eight foul shots, then a colored disk could be used to simulate a foul shot. A green disk could represent a made shot, and a red disk could represent a miss. You could put five green and three red disks in a container, mix them, and then choose one to represent a foul shot. If the color of the disk is green, then the shot is made. If the color of the disk is red, then the shot is missed. This procedure simulates one foul shot.

a. Using colored disks, describe how one at bat could be simulated for a baseball player who has a batting average of 0.300. Note that a batting average of 0.300 means the player gets a hit (on average) three times out of every ten times at bat. Be sure to state clearly what a color represents.

b. Using colored disks, describe how one at bat could be simulated for a player who has a batting average of 0.273. Note that a batting average of 0.273 means that on average, the player gets 273 hits out of 1,000 at bats.

Example 2: Using Random Number Tables

Why is using colored disks not practical for the situation described in Example 1(b)? Another way to carry out a simulation is to use a random number table, or a random number generator. In a random number table, the digits 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9 occur equally often in the long run. Pages and pages of random numbers can be found online.

For example, here are three lines of random numbers. The space after every five digits is only for ease of reading. Ignore the spaces when using the table.

25256	65205	72597	00562	12683	90674	78923	96568	32177	33855
76635	92290	88864	72794	14333	79019	05943	77510	74051	87238
07895	86481	94036	12749	24005	80718	13144	66934	54730	77140

To use the random number table to simulate an at bat for the 0.273 hitter in Example 1(b), you could use a three-digit number to represent one at bat. The three-digit numbers 000–272 could represent a hit, and the three-digit numbers 273–999 could represent a non-hit. Using the random numbers above and starting at the beginning of the first line, the first three-digit random number is 252, which is between 000 and 272, so that simulated at bat is a hit. The next three-digit random number is 566, which is a non-hit.

Continuing on the first line of the random numbers above, what would the hit/non-hit outcomes be for the next six at bats? Be sure to state the random number and whether it simulates a hit or non-hit.

Example 3: Baseball Player

A batter typically gets to bat four times in a ball game. Consider the 0.273 hitter from the previous example. Use the following steps (and the random numbers shown above) to estimate that player's probability of getting at least three hits (three or four) in four times at bat.

a. Describe what one trial is for this problem.

b. Describe when a trial is called a success and when it is called a failure.

c. Simulate 12 trials. (Continue to work as a class, or let students work with a partner.)

d. Use the results of the simulation to estimate the probability that a 0.273 hitter gets three or four hits in four times at bat. Compare your estimate with other groups.

Example 4: Birth Month

In a group of more than 12 people, is it likely that at least two people, maybe more, will have the same birth month? Why? Try it in your class.

Now, suppose that the same question is asked for a group of only seven people. Are you likely to find some groups of seven people in which there is a match but other groups in which all seven people have different birth months? In the following exercises, you will estimate the probability that at least two people in a group of seven were born in the same month.

Exercises 1–4

1. What might be a good way to generate outcomes for the birth month problem—using coins, number cubes, cards, spinners, colored disks, or random numbers?

2. How would you simulate one trial of seven birth months?

3. How is a success determined for your simulation?

4. How is the simulated estimate determined for the probability that a least two in a group of seven people were born in the same month?

Lesson Summary

To design a simulation:

- Identify the possible outcomes, and decide how to simulate them, using coins, number cubes, cards, spinners, colored disks, or random numbers.
- Specify what a trial for the simulation will look like and what a success and a failure would mean.
- Make sure you carry out enough trials to ensure that the estimated probability gets closer to the actual probability as you do more trials. There is no need for a specific number of trials at this time; however, you want to make sure to carry out enough trials so that the relative frequencies level off.

Lesson 12: Applying Probability to Make Informed Decisions

Example 1: Number Cube

Your teacher gives you a number cube with numbers 1–6 on its faces. You have never seen that particular cube before. You are asked to state a theoretical probability model for rolling it once. A probability model consists of the list of possible outcomes (the sample space) and the theoretical probabilities associated with each of the outcomes. You say that the probability model might assign a probability of $\frac{1}{6}$ to each of the possible outcomes, but because you have never seen this particular cube before, you would like to roll it a few times. (Maybe it is a trick cube.) Suppose your teacher allows you to roll it 500 times, and you get the following results:

Outcome	1	2	3	4	5	6
Frequency	77	92	75	90	76	90

Exercises 1–2

1. If the equally likely model was correct, about how many of each outcome would you expect to see if the cube is rolled 500 times?

2. Based on the data from the 500 rolls, how often were odd numbers observed? How often were even numbers observed?

Exercises 3–4

There are three popular brands of mixed nuts. Your teacher loves cashews, and in his experience of having purchased these brands, he suggests that not all brands have the same percentage of cashews. One has around 20% cashews, one has 25%, and one has 35%.

Your teacher has bags labeled A, B, and C representing the three brands. The bags contain red beads representing cashews and brown beads representing other types of nuts. One bag contains 20% red beads, another 25% red beads, and the third has 35% red beads. You are to determine which bag contains which percentage of cashews. You cannot just open the bags and count the beads.

3. Work as a class to design a simulation. You need to agree on what an outcome is, what a trial is, what a success is, and how to calculate the estimated probability of getting a cashew. Base your estimate on 50 trials.

4. Your teacher will give your group one of the bags labeled A, B, or C. Using your plan from part (a), collect your data. Do you think you have the 20%, 25%, or 35% cashews bag? Explain.