Lesson 10: Using Simulation to Estimate a Probability

Student Outcomes

- Students learn simulation as a method for estimating probabilities that can be used for problems in which it is difficult to collect data by experimentation or by developing theoretical probability models.
- Students learn how to perform simulations to estimate probabilities.
- Students use various devices to perform simulations (e.g., coin, number cube, cards).
- Students compare estimated probabilities from simulations to theoretical probabilities.

Lesson Overview

Simulation uses devices such as coins, number cubes, and cards to generate outcomes that represent real outcomes. Students may find it difficult to make the connection between device outcomes and the real outcomes of the experiment.

This lesson eases students into simulation by defining the real outcomes of an experiment and specifying a device to simulate the outcome. Then we will carefully lead students to define how an outcome of the device represents the real outcome, define a trial for the simulation, and identify what is meant by a trial resulting in a "success" or "failure." Be sure that students see how a device that may have many outcomes can be used to simulate a situation that has only two outcomes. For example, how a number cube can be used to represent a boy birth (e.g., even outcome, prime number outcome, or any three of its digits.)

Classwork

In previous lessons, you estimated probabilities of events by collecting data empirically or by establishing a theoretical probability model. There are real problems for which those methods may be difficult or not practical to use. Simulation is a procedure that will allow you to answer questions about real problems by running experiments that closely resemble the real situation.

It is often important to know the probabilities of real-life events that may not have known theoretical probabilities. Scientists, engineers, and mathematicians design simulations to answer questions that involve topics such as diseases, water flow, climate changes, or functions of an engine. Results from the simulations are used to estimate probabilities that help researchers understand problems and provide possible solutions to these problems.

Example 1 (10 minutes): Families

This first example begins with an equally likely model that simulates boy and girl births in a family of three children. Note that in human populations, the probabilities of a boy birth and of a girl birth are not actually equal, but we treat them as equal here. The example uses a coin in the simulation.

MP.2

There are five steps in the simulation:

- The first is to define the basic outcome of the real experiment, e.g., a birth.
- The second is to choose a device and define which possible outcomes of the device will represent an outcome of the real experiment, (e.g., toss of a coin, head represents boy; roll of a number cube, prime number (P) represents boy; choice of a card, black card represents boy.)
- The third is to define what is meant by a trial in the simulation that represents an outcome in the real experiment (e.g., three tosses of the coin represents three births; three rolls of a number cube represents three births; three cards chosen with replacement represents three births.)
- The fourth is to define what is meant by a success in the performance of a trial (e.g., using a coin, HHT represents exactly two boys in a family of three children; using a number cube, NPP represents exactly two boys in a family of three children; using cards, BRB represents exactly two boys in a family of three children.) Be sure that your students realize that in using a coin, HHT, HTH, and THH all represent exactly two boys in a family of three children whereas HHH is the only way to represent three boys in a family of three children.
- The fifth step is to perform n trials (the more the better), count the number of successes in the n trials, and divide the number of successes by n, which produces the estimate of the probability based on the simulation.

You may find it useful to reiterate the five steps for every problem in this lesson so that your students gain complete understanding of the simulation procedure.

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.

Pose the following questions to the class one at a time and allow for multiple responses:

- How could a number cube be used to simulate getting a boy or a girl birth?
  - *An even-number outcome represents boy, and an odd-number outcome represents girl; a prime-number outcome represents boy, and a non-prime outcome represents girl; or, any three-number cube digits can be chosen to represent boy with the rest of them, girl.*

- How could a deck of cards be used to simulate getting a boy or a girl birth?
  - *The most natural one would be black card for one gender, and red for the other.*

Exercises 1–2 (5 minutes)

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.

Let students work with a partner. Then discuss and confirm answers as a class.

1. **Suppose a prime number (P) result of a rolled number cube 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?**

   The outcomes are $2, 3, 5$ for a boy birth and $1, 4, 6$ for a girl birth. The boy and girl births are thereby 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.**

   The key response has to include the drawing of three cards with replacement. If a card is not replaced and the deck shuffled before the next card is drawn, then the probabilities of the genders have changed (ever so slightly, but they are not $50-50$ from draw to draw). Simulating the genders of three children requires three cards to be drawn with replacement.

Example 2 (5 minutes)

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 $\frac{28}{100} = 0.28$.

This example describes what is meant by a trial, a success, and how to estimate the probability of the desired event (i.e., that a family has three boys or three girls). It uses the coin device, 100 trials, in which 28 of them were either HHH or TTT. Hence the estimated probability that a family of three children has three boys or three girls is $\frac{28}{100}$. Then ask:

- What is the estimated probability that the three children are not all the same gender?
  - $1 - 0.28 = 0.72.$

Exercises 3–5 (15 minutes)

Let students continue to work with their partners on Exercises 3–5. Discuss the answer to Exercise 5 (c) as a class.

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

T represents a girl. Students should go through the list, count the number of times that they find HHT, HTH, or THH and divide that number of successes by 50. They should find the simulated probability to be $\frac{16}{50} = 0.32$. If time permits, ask students, "Based on what you found, is it likely that a family with three children will have exactly one girl? From your own experiences, how many families do you know who have three children and exactly one girl?"

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.

Answers will vary. For example, they could identify a girl birth as $1, 2, 3$ outcome on one roll of the number cube, and roll the number cube three times to simulate three children (one trial). They need to list their 50 trials. Note that an outcome of 412 would represent two girls, 123 would represent three girls, and 366 would represent one girl, as would 636 and 663. Be sure that they are clear about how to do all five steps of the simulation process.

MP.4

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).

The sample space is: BBB, BBG, BGB, GBB, BGG, GBG, GGB, GGG.

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.

Each is equally likely, so the theoretical probability of getting exactly one girl is $\frac{3}{8} = 0.375$ (BBG, BGB, GBB.)

c. Compare it to the estimated probabilities found in parts (a) and (b) above.

Answers will vary; the estimated probabilities from the first two parts of this exercise should be around 0.375. If not, you may suggest that they conduct more trials.

Example 3 (5 minutes): Basketball Player

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

Present the beginning of the example. This example has a real outcome probability of 0.75. Ask students:

- What device could be used to generate a probability of 3 out of 4?
  - *If they have studied Platonic solids, they may suggest a tetrahedron. Note that the outcome would be the result that is face down.*

Continue reading through the text of the example as a class. Explain how a number cube could be used in which a $1, 2,$ or 3 would represent a hit, the number 4 would represent a miss, and 5 and 6 would be ignored. The simulation is designed to fit the probabilities of 25% and 75%. Outcomes of $1, 2, 3,$ and 4 provide those probabilities. Adding the outcomes of 5 or 6 would change the probabilities; therefore, they are ignored (or simply not counted as outcomes).

Ask students:

- Is there another way to assign the numbers on the cube to the outcomes?
  - *Yes, 1 could represent a miss and $2, 3, 4$ could represent a hit.*

MP.5

50 trials of six numbers each are shown. Students are to estimate the probability that the player makes five or six of the six shots she takes. So, they should count how many of the trials have five or six of the numbers $1, 2, 3$ in them as successes. They should find the estimated probability to be $\frac{27}{50} = 0.54$.

As an aside, the theoretical probability is calculated by considering the possible outcomes for six foul shots that have either 6 successes (SSSSSS) or 5 successes (FSSSSS, SFSSSS, SSFSSS, SSSFSS, SSSSFS, and SSSSSF). The outcome that consists of six successes has probability $(0.75)^6$ and each of the 6 outcomes with five successes has probability $(0.75)^5(0.25)$. Based on the binominal probability distribution, the theoretical probability is $6(0.75)^5(0.25) + (0.75)^6$, or approximately 0.5339.

Closing (5 minutes)

> **Lesson Summary**
>
> **In previous lessons, you estimated probabilities by collecting data and found theoretical probabilities by creating a model. In this lesson you used simulation to estimate probabilities in real problems and in situations for which empirical or theoretical procedures are not easily calculated.**
>
> **Simulation is a method that uses an artificial process (like tossing a coin or rolling a number cube) to represent the outcomes of a real process that provides information about the probability of events. In several cases, simulations are needed to both understand the process as well as provide estimated probabilities.**

Exit Ticket (5–10 minutes)

Name ____ Date____________________

Lesson 10: Using Simulation to Estimate a Probability

Exit Ticket

1. Nathan is your school's star soccer player. When he takes a shot on goal, he scores half of the time on average. Suppose that he takes six shots in a game. To estimate probabilities of the number of goals Nathan makes, use simulation with a number cube. One roll of a number cube represents one shot.
   a. Specify what outcome of a number cube you want to represent a goal scored by Nathan in one shot.

   b. For this problem, what represents a trial of taking six shots?

   c. Perform and list the results of ten trials of this simulation.

   d. Identify the number of goals Nathan made in each of the ten trials you did in part (c).

   e. Based on your ten trials, what is your estimate of the probability that Nathan scores three goals if he takes six shots in a game?

2. Suppose that Pat scores 40% of the shots he takes in a soccer game. If he takes six shots in a game, what would one simulated trial look like using a number cube in your simulation?

Exit Ticket Sample Solutions

1. **Nathan is your school's star soccer player. When he takes a shot on goal, he scores half of the time on average. Suppose that he takes six shots in a game. To estimate probabilities of the number of goals Nathan makes, use simulation with a number cube. One roll of a number cube represents one shot.**

   a. **Specify what outcome of a number cube you want to represent a goal scored by Nathan in one shot.**

      Answers will vary; students need to determine which three numbers on the number cube represent scoring a goal.

   b. **For this problem, what represents a trial of taking six shots?**

      Rolling the cube six times represents taking six shots on goal or 1 simulated trial.

   c. **Perform and list the results of ten trials of this simulation.**

      Answers will vary; students working in pairs works well for these problems. Performing only ten trials is a function of time. Ideally many more trials should be done. If there is time, have your class pool their results.

   d. **Identify the number of goals Nathan made in each of the ten trials you did in part (c).**

      Answers will vary.

   e. **Based on your ten trials, what is your estimate of the probability that Nathan scores three goals if he takes six shots in a game?**

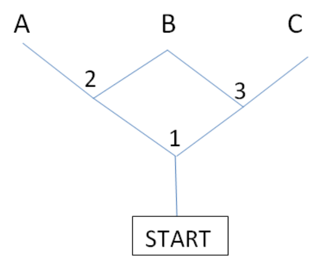
      Answers will vary; the probability of scoring per shot is $\frac{1}{2}$.

2. **Suppose that Pat scores 40% of the shots he takes in a soccer game. If he takes six shots in a game, what would one simulated trial look like using a number cube in your simulation?**

   Students need to realize that 40% ***is*** 2 ***out of*** 5***. In order to use the number cube as the device,*** 1 ***and*** 2 ***could represent goals, while*** $3, 4,$ ***and*** 5 ***could represent missed shots, and*** 6 ***is ignored. Rolling the number cube six times creates*** 1 ***simulated trial.***

Problem Set Sample Solutions

1. A mouse is placed at the start of the maze shown below. If it reaches station B, it is given a reward. At each point where the mouse has to decide which direction to go, assume that it is equally likely to go in either direction. At each decision point $1, 2, 3$, it must decide whether to go left (L) or right (R). It cannot go backwards.



a. Create a theoretical model of probabilities for the mouse to arrive at terminal points A, B, and C.

i. List the possible paths of a sample space for the paths the mouse can take. For example, if the mouse goes left at decision point 1, and then right at decision point 2, then the path would be denoted LR.

The possible paths in the sample space are {LL, LR, RL, RR}.

ii. Are the paths in your sample space equally likely? Explain.

Each of these outcomes has an equal probability of $\frac{1}{4}$, since at each decision point there are only two possible choices, which are equally likely.

iii. What are the theoretical probabilities that a mouse reaches terminal points A, B, and C? Explain.

The probability of reaching terminal point A is $\frac{1}{4}$, since it is accomplished by path LL. Similarly, reaching terminal point C is $\frac{1}{4}$, since it is found by path RR. However, reaching terminal point B is $\frac{1}{2}$, since it is reached via LR or RL.

b. Based on the following set of simulated paths, estimate the probabilities that the mouse arrives at points A, B, and C.

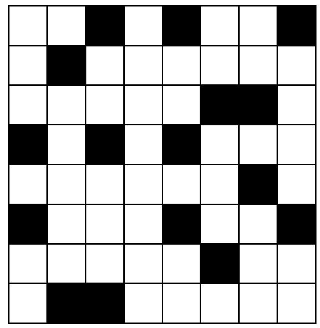
RR	RR	RL	LL	LR	RL	LR	LL	LR	RR
LR	RL	LR	RR	RL	LR	RR	LL	RL	RL
LL	LR	LR	LL	RR	RR	RL	LL	RR	LR
RR	LR	RR	LR	LR	LL	LR	RL	RL	LL

Students need to go through the list and count the number of paths that go to A, B, and C. They should find the estimated probabilities to be $\frac{8}{40} = 0.2$ for A, $\frac{22}{40} = 0.55$ for B, and $\frac{10}{40} = 0.25$ for C.

c. How do the simulated probabilities in part (b) compare to the theoretical probabilities of part (a)?

The probabilities are reasonably close for parts (a) and (b). Students should realize that probabilities based on taking 400 trials should be closer than those based on 40, but that the probabilities based on 40 are in the ballpark.

2. **Suppose that a dartboard is made up of the 8×8 grid of squares shown below. Also, suppose that when a dart is thrown, it is equally likely to land on any one of the 64 squares. A point is won if the dart lands on one of the 16 black squares. Zero points are earned if the dart lands in a white square.**



   a. **For one throw of a dart, what is the probability of winning a point? Note that a point is won if the dart lands on a black square.**

   The probability of winning a point is $\frac{16}{64} = 0.25$.

   b. **Lin wants to use a number cube to simulate the result of one dart. She suggests that 1 on the number cube could represent a win. Getting $2, 3$, or 4 could represent no point scored. She says that she would ignore getting a 5 or 6. Is Lin's suggestion for a simulation appropriate? Explain why you would use it *or,* if not, how you would change it.**

   Lin correctly suggests that to simulate the result of one throw, a number cube could be used with the 1 representing a hit, $2, 3, 4$ representing a missed throw, while ignoring 5 and 6. (As an aside, a tetrahedron could be used by using the side facing down as the result.)

   c. **Suppose a game consists of throwing a dart three times. A trial consists of three rolls of the number cube. Based on Lin's suggestion in part (b) and the following simulated rolls, estimate the probability of scoring two points in three darts.**

324	332	411	322	124
224	221	241	111	223
321	332	112	433	412
443	322	424	412	433
144	322	421	414	111
242	244	222	331	224
113	223	333	414	212
431	233	314	212	241
421	222	222	112	113
212	413	341	442	324

   The probability of scoring two points in three darts is $\frac{5}{50} = 0.1$. *(Students need to count the number of trials that contain exactly two 1's.)*

d. The theoretical probability model for winning 0, 1, 2, and 3 points in three throws of the dart as described in this problem is

i. winning 0 points has a probability of 0.42;

ii. winning 1 point has a probability of 0.42;

iii. winning 2 points has a probability of 0.14;

iv. winning 3 points has a probability of 0.02.

Use the simulated rolls in part (c) to build a model of winning 0, 1, 2, and 3 points, and compare it to the theoretical model.

To find the estimated probability of 0 points, count the number of trials that have no 1's in them ($\frac{23}{50} = 0.46$).

To find the estimated probability of 1 point, count the number of trials that have one 1 in them ($\frac{20}{50} = 0.4$).

From part (c), the estimated probability of 2 points is 0.1.

To find the estimated probability of 3 points, count the number of trials that have three 1's in them ($\frac{2}{50} = 0.04$.)

The theoretical and simulated probabilities are reasonably close.

	0	1	2	3
Theoretical	0.42	0.42	0.14	0.02
Simulated	0.46	0.40	0.10	0.04