Building Concepts: Law of Large Numbers

Lesson Overview

In this TI-Nspire lesson, students investigate how the relative frequency of an outcome approaches the actual probability of that outcome as the number of repetitions gets larger and larger (the law of large numbers).

Over the long run, the relative frequencies of outcomes of chance processes stabilize as the sample size gets larger and can be used to estimate probabilities.

Learning Goals

1. Identify the set of possible outcomes from a chance event and estimate the probabilities of those outcomes from repeating the chance process a large number of times;
2. differentiate between frequency and relative frequency when referring to the outcomes of a chance process;
3. recognize that the relative frequency of an outcome is likely to be close to the actual probability of that outcome as the number of repetitions gets larger and larger (the law of large numbers);
4. recognize that relative frequencies of particular outcomes after just a few observations are more likely to deviate substantially from expected values than are relative frequencies based on more observations.

Prerequisite Knowledge

Law of Large Numbers is the sixteenth lesson in a series of lessons that explore the concepts of statistics and probability. This lesson builds on the concepts of the previous lessons. Prior to working on this lesson students should have completed *What is Probability?* and *Probability, Diagrams, and Tables*. Students should understand:

- how to read a bar graph;
- how to identify a sample space;
- the concept of theoretical probability.

Vocabulary

- **distribution:** the way in which data is spread out
- **variability:** the spread in a group of data
- **sample space:** the collection of all possible individual outcomes
- **outcome:** a possible result of a probability experiment
- **frequency:** the number of times an event occurs in a chance experiment.
- **relative frequency:** the number of times an event occurs during experimental trials, divided by the total number of trials conducted

Lesson Pacing

This lesson should take 50–90 minutes to complete with students, though you may choose to extend, as needed.

Lesson Materials

- Compatible TI Technologies:

  TI-Nspire CX Handhelds, TI-Nspire Apps for iPad®, TI-Nspire Software
- Law of Large Numbers_Student.pdf
- Law of Large Numbers_Student.doc
- Law of Large Numbers.tns
- Law of Large Numbers_Teacher Notes
- To download the TI-Nspire activity (TNS file) and Student Activity sheet, go to http://education.ti.com/go/buildingconcepts.

Class Instruction Key

The following question types are included throughout the lesson to assist you in guiding students in their exploration of the concept:

Class Discussion: Use these questions to help students communicate their understanding of the lesson. Encourage students to refer to the TNS activity as they explain their reasoning. Have students listen to your instructions. Look for student answers to reflect an understanding of the concept. Listen for opportunities to address understanding or misconceptions in student answers.

Student Activity: Have students break into small groups and work together to find answers to the student activity questions. Observe students as they work and guide them in addressing the learning goals of each lesson. Have students record their answers on their student activity sheet. Once students have finished, have groups discuss and/or present their findings. The student activity sheet can also be completed as a larger group activity, depending on the technology available in the classroom.

Deeper Dive: These questions are provided for additional student practice and to facilitate a deeper understanding and exploration of the content. Encourage students to explain what they are doing and to share their reasoning.

Mathematical Background

A *probability model* provides a probability for each possible distinct outcome for a chance process where the total probability over all such outcomes is 1. The collection of all possible individual outcomes is known as the *sample space* for the model. In this lesson students compare the distribution of the entire sample space to distributions of randomly generated outcomes from that sample space. Generating single outcomes from a given sample space and observing how the shape of the sample distribution (represented as a bar graph) takes form as the number of observations increases shows an application of the law of large numbers, i.e., the distributions of sample outcomes from a random process stabilize as the sample size gets larger.

In Activity12, *What is Probability?*, students repeatedly drew chips at random from a bag with an unknown proportion of blue and white chips. They estimated the probability of drawing a blue chip from a bag of chips by considering the long run relative frequency of obtaining a blue chip when repeatedly drawing, chips one at a time from the bag (with replacement). This activity extends on that idea and builds a frequency distribution (or counts) of several outcomes, that are not all equally likely. Because the outcomes are not equally likely, the distributions are not uniform, which connects to the work in Activity 14, *Unequally Likely Outcomes*.

In an effort to compare the effect of different sample sizes on the distribution of outcomes, the outcomes are represented as relative frequencies. The difference between a frequency distribution and a relative frequency distribution is a key distinction that must be made in order for students to understand the effect of sample size on the sample distribution. As the sample size increases (more and more counts are added), the relative frequency for each outcome approaches the theoretical probability of that outcome.

Part 1, Page 1.3

Focus: The probabilities of chance outcomes can be estimated by repeating the chance process a large number of times.

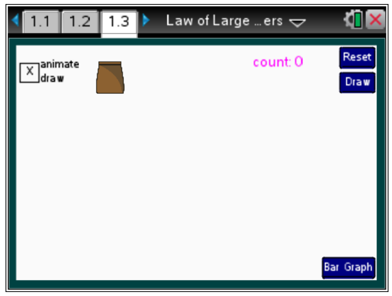
On page 1.3, students select the bag to draw one chip from the bag, with replacement.

animate draw toggles the animation on or off.

Bar Graph shows a bar graph of the counts for each color jellybean.

Draw shows one chip drawn from the bag.

Draw 5 displays after bar graph has been generated and shows the results of drawing one chip five times.



TI-Nspire Technology Tips

menu accesses page options.

tab toggles through options on the page.

enter toggles animation and controls **Draw**.

ctrl del resets the page.

Class Discussion

The following questions help establish the concept that small samples have a lot of variability while laying the foundation for the terms *sample space* and *frequency distribution*.

The bag on page 1.3 contains a variety of different color jellybeans. Selecting the bag will draw a jellybean from the bag and plot the jellybeans according to color. After each draw, the jellybean is returned to the bag.

Have students…	Look for/Listen for…
• ***Draw until you have ten jellybeans. How many different jellybean colors do you think are in the bag? Explain your reasoning.***	Answers will vary. Some students may say they can't tell without drawing more jellybeans. Others might respond, "I have four colors and six jellybeans are yellow."
• ***Predict the number of each jellybean color you will have if you generate 10 more jellybeans and add them to the distribution.***	Answers will vary. Students may base their predictions on the distribution they have already generated, indicating "more of the same," so the predictions might be double their results from the first ten draws.
• ***Select* animate/draw*, then* Draw *to draw 10 more jellybeans. Were you surprised by the resulting distribution? Explain why or why not.***	Answers will vary. Most students should be surprised because with only 10 jellybeans, the distribution of results will often be very different from the distribution of 20 jellybeans.

Class Discussion (continued)

• ***Draw 30 more jellybeans for a total of 50 jellybeans. Use your results to make a prediction about the proportion of jellybeans in the bag that are each color.***	Answers will vary. In one example, 52% were yellow, 26% purple, 12% blue, 4% orange, and 6% green.
Draw another 50 jellybeans for a total of 100 jellybeans.	
• ***How do your results compare to those from the question above?***	Answers will vary. One example gave 42% purple, 30% yellow, 14% orange, 8% blue and 6% green.
• ***Compare your results to other students. What are the similarities and differences among your distributions of the colors of jellybeans in the bag?***	Answers will vary. Students should note that purple and yellow are the most frequent, and the numbers of the other colors vary but are not as frequent as purple and yellow.
• ***Sally says she can predict the percentage of yellow jellybeans in the bag from looking at the distribution of 50 jellybeans she drew from the bag. What would you say to Sally?***	Answers will vary. Her prediction will probably be off a bit because the number of yellow jellybeans varied so much in our different samples.
• ***Draw another 50 jelly beans and compare your results to others in class.***	Answers will vary, but the distributions should begin to look more similar.

Student Activity Questions—Activity 1

1. **The set of all possible outcomes from a chance experiment is called the sample space for that experiment.**

   a. **What is the sample space for drawing a jellybean from the bag?**

   Answer: The five colors: blue, green, orange, purple, and yellow.

   b. **Are the outcomes for each color equally likely? Why or why not?**

   Answer: No, because the chance of getting a yellow or a purple appears greater than the chance of getting one of the other three colors.

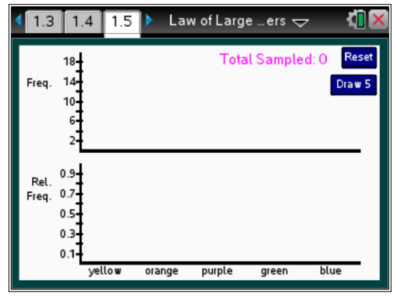
   c. **Reset and select Bar Graph. The graph shows a *frequency* distribution. Draw jellybeans until you have 50. Make a sketch of your results. Which color had the greatest frequency? The least?**

   Answers will vary: In one example, yellow jellybeans occurred with the greatest frequency and blue with the least.

   d. **Reset and draw another 50 jellybeans. How does this frequency distribution compare to the one from question c above?**

   Answers will vary, but typically the distributions will not be the same. In a second set of 50 jellybeans, the yellow and purple jellybeans were the same and the greatest frequency and green the least.

Part 2, Page 1.5



Focus: The *frequency* of an outcome is the number of times an outcome occurs while the *relative frequency* of the outcome is the number of times the outcome occurs divided by the total number of repetitions of the chance process.

Page 1.5 shows the frequency (the number of each color jellybean for five draw) on the top plot, and the relative frequency of each color jellybean for the five draws on the bottom plot.

Draw 5 shows the frequency and relative frequency of each color after drawing five jellybeans one at a time and replacing each after it has been drawn.

Student Activity Questions—Activity 2

The following questions make explicit the difference between *count* or *frequency* and *relative frequency* and then connect these ideas to probability.

1. **On page 1.5 draw 10 jellybeans.**

   a. **Explain what the top and bottom graphs represent. Use the distribution of the jellybean colors to support your answer. Note that moving the cursor over a bar displays information about what the bar represents.**

   Answer: The top graph shows the count for each color jellybean, so in my example I had 4 yellow and 4 purple jellybeans, 1 orange and 1 green. The bottom shows the proportion of the ten jellybeans for each color, so for my example yellow and purple would be 0.4 ($\frac{4}{10}$); orange and green would be 0.1 ($\frac{1}{10}$).

   b. **The proportion or percent of jellybeans that are each color is called the *relative frequency* for that color. Explain how the relative frequency is different from just the frequency.**

   Answer: The *frequency* is just the count for how many times a color was drawn. The *relative frequency* is how many times a certain color jellybean occurred divided by the total number of jellybeans drawn.

   c. **Suppose Adam had a distribution of the colors with 50 jellybeans, 18 of which are blue and Bethany had a distribution of the colors with 75 jellybeans, 25 of which are blue. Whose distribution had the greater relative frequency of blue jellybeans?**

   Answer: Since $\frac{18}{50} = 0.36$ is greater than $\frac{25}{75} = 0.33$, Adam's distribution had the greater relative frequency of blue jellybeans.

Student Activity Questions—Activity 2 (continued)

d. **Think about drawing 50 and then drawing 100 jellybeans. Why is the concept of *relative frequency* important?**

Answer: If you get 20 yellow jellybeans you wouldn't know whether it was out of 50 or out of 100. The relative frequency makes sure that the total number is considered in thinking about the proportion of yellow jellybeans, not just the absolute number, i.e., $\frac{20}{50}$ would be 0.4, while 20 out of 100 would be 0.2.

2. **Remember experimental probability from Activity 12, *What is Probability?*. You can use the long-run relative frequency of an outcome to estimate the probability of that outcome. (Note this is sometimes written as *P*(outcome).)**

a. **Use your results to estimate the probability of getting a purple jellybean.**

Answers will vary. For the example in 1.a., the probability of getting a purple jellybean would be $\frac{4}{10}$ or 0.4.

b. **Draw 40 more jellybeans for a total of 50. How did your estimated probability for getting a purple jellybean from the question above change?**

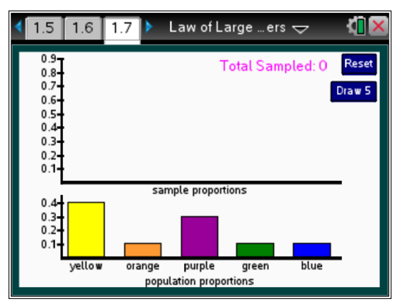
Answers will vary. In one example, $P(\text{purple}) = 0.24$ or $\frac{6}{25}$.

c. **In a sample of 50 jellybeans, the estimate for $P(\text{blue}) = 0.15$, $P(\text{orange}) = 0.09$, $P(\text{green}) = 0.12$, $P(\text{purple}) = 0.35$. What is an estimate for *P*(yellow) and how do you know?**

Answer: An estimate for $P(\text{yellow}) = 1-(0.15+0.09+0.12+0.35) = 0.29$. The sum of all of the estimated probabilities has to equal 1.00.

Part 3, Pages 1.7

Focus: The relative frequency of an outcome is likely to be close to the actual probability of that outcome as the number of repetitions gets larger and larger (the law of large numbers).



On page 1.7, the bottom line shows the relative frequency of the colors in the whole bag of jellybeans.

Draw 5 shows the relative frequency of each color after drawing five jellybeans, one at a time, and replacing each after it has been drawn.

Class Discussion

These questions explore the relationship between a long-run relative frequency distribution and the actual population distribution.

On page 1.7, the bottom graph represents the distribution of the actual proportions of the five colors of jellybeans. Moving over the bars displays the percent that color is of all the colors.

Have students…	**Look for/Listen for…**
• ***If you draw a jellybean from the bag, what is the probability of getting each jellybean color?***	Answer: $P(\text{blue}) = P(\text{green}) = P(\text{orange}) = 0.1$, $P(\text{purple}) = 0.3$, $P(\text{yellow}) = 0.4$
• ***Draw five jellybeans from the bag. Explain what you see on the two graphs.***	Answers may vary. The bottom axis still shows the proportion of each color jellybean in the bag. The top axis shows the relative frequency of each color when you drew a sample of 5 jellybeans. In my sample, 60% were yellow, 20% were blue, and 20% were purple.
• **Draw a total of 200 jellybeans. What is the estimated probability of getting a blue jellybean?**	Answers may vary. In one sample, the estimated probability of getting a blue jellybean was 0.09.
• ***What should be the sum of all of the estimated probabilities? Explain your reasoning. Use the results of the question above to check your answer.***	Answer: The sum should be 1.00.

Student Activity Questions—Activity 3

1. Seth wrote out the list below of the things he learned in this lesson. What would you say to Seth? Use an example from the TNS activity to support your advice.

a. "A sample of about 20 will give you a fairly good idea about the distribution of the outcomes of an experiment (like the colors of jellybeans when you drawing them out of a bag)."

Answer: This is not correct because after I drew 20 jellybeans, sometimes I did not even have one jellybean of each color. You need more draws to get a good idea of the proportion of each color. Since there are so many different colors, it seems to take almost 200 draws to get a fairly good idea of the proportion of each color.

Student Activity Questions—Activity 3 (continued)

b. "The outcomes of a chance process do not all have to have the same chance of occurring (be equally likely)."

Answer: Seth is right. In this case the proportions of the colors were not all the same; the chance of getting yellow was different from the chance of getting green.

c. "The frequency of an outcome and the relative frequency of the same outcome are the same thing."

Answer: Seth is not right. The frequency is just how many times the outcome occurred; the relative frequency is the number of times the outcome occurs, divided by the total number of trials conducted. So, if you drew 20 jellybeans and got 8 yellow, the 8 is the frequency but the relative frequency would be $\frac{8}{20}$ or 0.4.

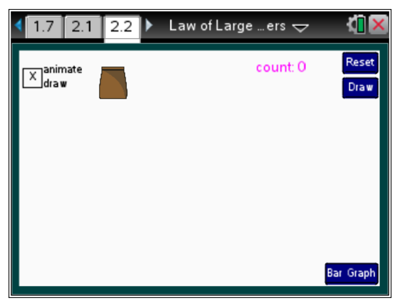
d. "After drawing lots and lots of jellybeans, the relative frequency of each color jellybean got closer and closer to the actual proportion of each color in the bag."

Answer: Seth is right. When you drew 200 jellybeans, the relative frequency distribution of the colors was almost the same as the frequency distribution of the colors.

Part 4, Page 2.2

Focus: Small samples can be misleading.

Page 2.2 has the same functionality as page 1.3.



Student Activity Questions—Activity 4

1. **Draw 10 jellybeans and enter the observed frequency of each jellybean color in the first row of the table. (Note: Each repetition consists of exactly 10 jellybeans.) Reset and draw a new set of ten jellybeans. Record the frequency of each color in the table. Fill in the table using the same process.**

repetition	yellow	orange	purple	green	blue
1	6	0	2	0	2
2	4	3	2	1	0
3	4	0	2	1	3
4	3	1	5	1	0
5	2	0	6	1	1

Student Activity Questions—Activity 4 (continued)

a. **Which of the samples you drew might be misleading about the proportion of colors for the jellybeans? Explain your reasoning.**

Answers will vary. Using the example above, some may note that in the first sample, you only had three colors; and in all of the others you only had four.

b. **Based on your results in the table, do you think this is the same bag of jellybeans used on page 1.3? Why or why not?**

Answer: Students should notice that the outcomes for each color vary widely in the five trials and make it impossible to tell whether the jellybeans come from the same bag of jellybeans. For example, the first set of 10 had only three colors. Some may decide to accumulate all of each color, getting for example 6 out of 50 blue (12%), 4 out of 50 green (8%), 4 out of 50 orange (8%), 17 out of 50 purple (34%) and 19 out of 50 yellow (38%). The percentages are close to the previous bag (original 40% yellow, 30% purple and 10% for the other three colors) but there is still really not enough evidence to accept the fact that this is a new bag.

Deeper Dive — Pages 1.3, 1.5, and 1.7

Answer each of the questions below. Give an example to support your thinking.

- ***Explain how you would create a relative frequency distribution if you were given a frequency distribution.***

  Answer: You would add the counts or frequencies of each outcome to find a total. To find the relative frequency of each outcome, you would divide the frequency of the outcome by the total. For example, if you have 26 reds, 8 blues, 5 greens and 11 yellow, the total would be 50; the relative frequency for each color would be 0.52, 0.16, 0.1, and 0.22.

- ***Suppose you were given a relative frequency distribution. Can you create the frequency distribution? Why or why not?***

  Answer: You cannot go back to the frequency distribution from the relative frequency distribution unless you know what the original total was. For example, if the relative frequencies were: red: 0.2; blue: 0.3; green: 0.15; yellow: 0.35, you would not know whether the original total was 50 or 100 or 1000 because red could have come from $\frac{2}{10}=\frac{20}{100}=\frac{200}{1000}$ and so on. It would be the same for the other colors.

- ***How many jellybeans do you have to draw before the relative frequency distribution of the number of each color is exactly the same as the actual distribution of the colors?***

  Answer: It will almost never be exactly the same because the outcomes will always have some variability although when you have drawn 500 jellybeans, the differences in the two distributions of colors are likely very small.

Sample Assessment Items

After completing the lesson, students should be able to answer the following types of questions. If students understand the concepts involved in the lesson, they should be able to answer the following questions without using the TNS activity.

1. If you knew the sample space consisted of four outcomes, which of the following is true?

   a. The probability of each outcome is $\frac{1}{4}$. ***Answer: False***

   b. The sum of the probabilities of each outcome is 1. ***Answer: True***

   c. None of the probabilities of the outcomes is greater than $\frac{1}{2}$ ***Answer: False***

   d. The probability of one outcome is $\frac{1}{4}$. ***Answer: False***

2. Todd was practicing throwing darts. For each dart tossed, he scored the following points: 0, 15, 15, 30, 0, 45, 0, 15, 15, 30.

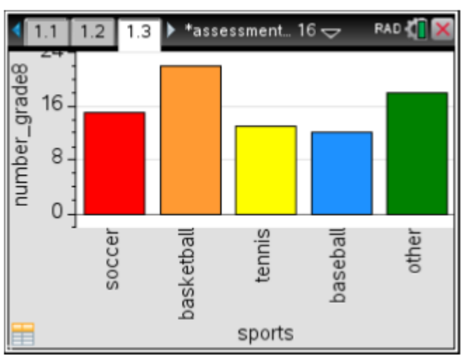
   a. Find Todd's relative frequency for scoring 0 points with a tossed dart.

   Answer: $\boldsymbol{\frac{3}{10}}$ ***or 0.3***

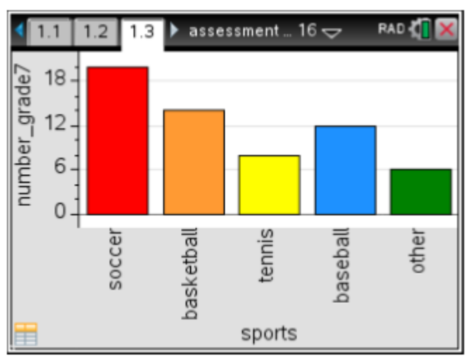
   b. Based on his practice tosses, which number of points is Todd most likely to earn with a dart toss?

   Answer: 15

3.



Eighth Grade Favorite Sports



Seventh Grade Favorite Sports

The results of a survey of seventh and eighth graders' favorite sports are displayed above.

   a. Find the probability that a randomly chosen seventh grader would prefer basketball.

   Answer: $\boldsymbol{\frac{14}{60}}$

b. Which grade has the greater relative frequency for soccer?

Answer: Eighth grade: $\frac{15}{80} \cong 0.188$ ***; Seventh grade:*** $\frac{20}{60} \cong 0.333$ ***, so seventh graders have the greatest relative frequency in preferring soccer.***

c. Shawana noted that the probability that a randomly chosen seventh grader would have baseball as a favorite sport would be the same as the probability that a randomly chosen eighth grader would have baseball as a favorite sport. Do you agree with Shawana? Why or why not?

Answer: Shawana is not correct because 80 eighth graders took the survey and 60 seventh graders. So $P(\text{baseball for eighth grader}) = \frac{12}{80}$***, but*** $P(\text{baseball for seventh grader}) = \frac{12}{60}$.

4. Benita and Jeff each surveyed some of the students in their eighth-grade homerooms to determine whether chicken or hamburgers should be served at the class picnic. The survey forms are shown below.

Benita's Survey
Homeroom: 8-A
Number of Students in Homeroom: 23

Student: Surveyed	Chicken	Hamburger
Adam	✓	
Carlene	✓	
Nancy	✓	
Hugh	✓	

Jeff's Survey
Homeroom: 8-B
Number of Students in Homeroom: 20

Student: Surveyed	Chicken	Hamburger
Becky		✓
Tanya	✓	
Joe	✓	
Ben		✓
Abby		✓
Linc	✓	
Marian		✓
Han		✓
Chris		✓
Tina		✓
Nate		✓
Darrell		✓

Benita reported that 100 percent of those in her survey wanted chicken. Jeff reported that 75 percent of those in his survey wanted hamburger. Which survey, Benita's or Jeff's, would probably be better to use when making the decision about what to serve?

Explain why that survey would be better.

NAEP, Grade 8, 2007

Answer: Jeff's survey because he asked more people. Benita surveyed only four out of 23 people and Jeff surveyed 12 out of 20, which will be more representative of the class.

Student Activity Solutions

In these activities you will use frequency and relative frequency to analyze distributions of jellybeans. After completing the activities, discuss and/or present your findings to the rest of the class.

Activity 1 [Page 1.3]

1. The set of all possible outcomes from a chance experiment is called the sample space for that experiment.

   a. What is the sample space for drawing a jellybean from the bag?

   Answer: The five colors: blue, green, orange, purple, and yellow.

   b. Are the outcomes for each color equally likely? Why or why not?

   Answer: No, because the chance of getting a yellow or a purple appears greater than the chance of getting one of the other three colors.

   c. Reset and select Bar Graph. The graph shows a *frequency* distribution. Draw jellybeans until you have 50. Make a sketch of your results. Which color had the greatest frequency? The least?

   Answers will vary: In one example, yellow jellybeans occurred with the greatest frequency and blue with the least.

   d. Reset and draw another 50 jellybeans. How does this frequency distribution compare to the one from question c above?

   Answers will vary, but typically the distributions will not be the same. In a second set of 50 jellybeans, the yellow and purple jellybeans were the same and the greatest frequency and green the least.

Activity 2 [Page 1.5]

1. On page 1.5 draw 10 jellybeans.

   a. Explain what the top and bottom graphs represent. Use the distribution of the jellybean colors to support your answer. Note that moving the cursor over a bar displays information about what the bar represents.

   Answer. The top graph shows the count for each color jellybean, so in my example I had 4 yellow and 4 purple jellybeans, 1 orange and 1 green. The bottom shows the proportion of the ten jellybeans for each color, so for my example yellow and purple would be 0.4 ($\frac{4}{10}$); orange and green would be 0.1 ($\frac{1}{10}$).

   b. The proportion or percent of jellybeans that are each color is called the *relative frequency* for that color. Explain how the relative frequency is different from just the frequency.

   Answer. The frequency is just the count for how many times a color was drawn. The relative frequency is how many times a certain color jellybean occurred divided by the total number of jellybeans drawn.

c. Suppose Adam had a distribution of the colors with 50 jellybeans, 18 of which are blue and Bethany had a distribution of the colors with 75 jellybeans, 25 of which are blue. Whose distribution had the greater relative frequency of blue jellybeans?

Answer. Since $\frac{18}{50}=0.36$ *which is greater than* $\frac{25}{75}=0.33$*, Adam's distribution had the greater relative frequency of blue jellybeans.*

d. Think about drawing 50 and then drawing 100 jellybeans. Why is the concept of "relative frequency" important?

Answer. If you get 20 yellow jellybeans you wouldn't know whether it was out of 50 or out of 100. The relative frequency makes sure that the total number is considered in thinking about the proportion of yellow jellybeans, not just the absolute number, i.e., $\frac{20}{50}$ *would be 0.4, while 20 out of 100 would be 0.2.*

2. Remember experimental probability from Activity 12, *What is Probability?* You can use the long run relative frequency of an outcome to estimate the probability of that outcome. (Note this is sometimes written as P(outcome).)

a. Use your results to estimate the probability of getting a purple jellybean.

Answers will vary. For the example in 1.a., the probability of getting a purple jellybean would be $\frac{4}{10}$ *or 0.4.*

b. Draw 40 more jellybeans for a total of 50. How did your estimated probability for getting a purple jellybean from the question above change?

Answers will vary. In one example, $P(\text{purple})=0.24$ or $\frac{6}{25}$.

c. In a sample of 50 jellybeans, the estimate for $P(\text{blue})=0.15$, $P(\text{orange})=0.09$, $P(\text{green})=0.12$, $P(\text{purple})=0.35$. What is an estimate for P(yellow) and how do you know?

Answer: An estimate for $P(\text{yellow})=1-(0.15+0.09+0.12+0.35)=0.29$*. The sum of all of the estimated probabilities has to equal 1.00.*

Activity 3 [Page 1.7]

1. Seth wrote out the list below of the things he learned in this lesson. What would you say to Seth? Use an example from the TNS activity to support your advice.

a. "A sample of about 20 will give you a fairly good idea about the distribution of the outcomes of an experiment (like the colors of jellybeans when you drawing them out of a bag)."

Answer: This is not correct because after I drew 20 jellybeans, sometimes I did not even have one jellybean of each color. You need more draws to get a good idea of the proportion of each color. Since there are so many different colors, it seems to take almost 200 draws to get a fairly good idea of the proportion of each color.

b. “The outcomes of a chance process do not all have to have the same chance of occurring (be equally likely).”

Answer: Seth is right. In this case the proportions of the colors were not all the same; the chance of getting yellow was different from the chance of getting green.

c. “The frequency of an outcome and the relative frequency of the same outcome are the same thing.”

Answer: Seth is not right. The frequency is just how many times the outcome occurred; the relative frequency is how many times it occurred divided by the total of all you looked at. So, if you drew 20 jellybeans and got 8 yellow, the 8 is the frequency but the relative frequency would be $\frac{8}{20}$ *or 0.4.*

d. “After drawing lots and lots of jellybeans, the relative frequency of each color jellybean got closer and closer to the actual proportion of each color in the bag.”

Answer. Seth is right. When you drew 200 jellybeans, the relative frequency distribution of the colors was almost the same as the frequency distribution of the colors.

Activity 4 [Page 2.2]

1. Draw 10 jellybeans and enter the observed frequency of each jellybean color in the first row of the table. (Note: Each repetition consists of exactly 10 jellybeans.) Reset and draw a new set of ten jellybeans. Record the frequency of each color in the table. Fill in the table using the same process.

repetition	yellow	orange	purple	green	blue
1	6	0	2	0	2
2	4	3	2	1	0
3	4	0	2	1	3
4	3	1	5	1	0
5	2	0	6	1	1

a. Which of the samples you drew might be misleading about the proportion of colors for the jellybeans? Explain your reasoning.

Answers will vary. Using the example above, some may note that in the first sample, you only had three colors; and in all of the others you only had four.

b. Based on your results in the table, do you think this is the same bag of jellybeans used on page 1.3? Why or why not?

Answer: Students should notice that the outcomes for each color vary widely in the five trials and make it impossible to tell whether the jellybeans come from the same bag of jellybeans. For example, the first set of 10 had only three colors. Some may decide to accumulate all of each color, getting for example 6 out of 50 blue (12%), 4 out of 50 green (8%), 4 out of 50 orange (8%), 17 out of 50 purple (34%) and 19 out of 50 yellow (38%). The percentages are close to the previous bag (original 40% yellow, 30%, purple and 10% for the other three colors) but there is still really not enough evidence to accept the fact that this is a new bag.