Designing and Conducting Simulations

ACTIVITY 40

Time Flies When You Are Having Fun

Lesson 40-1 Random Chance

My Notes

Learning Target:

- Determine whether an apparent treatment effect is too large to be due just to random chance.

SUGGESTED LEARNING STRATEGIES: Activating Prior Knowledge, KWL Chart, Role Play, Summarizing, Paraphrasing, Think Aloud, Debriefing

Jamie and Riley wanted to see whether the adage "Time flies when you are having fun!" could be demonstrated scientifically. They decided to conduct a study and recruited 14 classmates to be subjects. Jamie randomly selected seven subjects and assigned them to a group called "Fun." The rest were assigned to a group called "Not Fun." The "Fun" group was given a task of playing a video game enjoyed by all subjects, while the "Not Fun" group was assigned a task of copying code for a programming language with which none were familiar.

1. Is the study described an observational study or an experiment? Explain your reasoning.
   The study is an experiment since both groups have a treatment applied to them.

Each subject was asked to spend 30 minutes in a quiet room performing their assigned task with no time-keeping capability. At exactly 13.5 minutes into their task, subjects were interrupted and asked to estimate the number of minutes that had passed since the task began.

2. What are the variables in this study?
   The response (dependent) variable in this study is the perceived number of minutes that had passed. The explanatory variable is the treatment imposed, either performing a "fun" task or "not fun" task.

Common Core State Standards for Activity 40

HSS-IC.B.5 Use data from a randomized experiment to compare two treatments; use simulations to decide if differences between parameters are significant.

ACTIVITY 40

Investigative

Activity Standards Focus

In this activity students will build on their prior learning about simulations. Students will benefit if they already have a clear understanding of why simulations are done. Students will use simulations to determine statistical significance.

Lesson 40-1

PLAN

Pacing: 1 class period

Chunking the Lesson

#1–2 #3–4 #5–7

#8–9 #10–12

Check Your Understanding

Lesson Practice

TEACH

Bell-Ringer Activity

Have students give an example of two things that have a cause-and-effect relationship, explain the relationship, and describe the explanatory and response variables.

1–2 Shared Reading, Summarizing, Paraphrasing, KWL Chart Use reading strategies to be sure students understand the math terms. An *observation* involves studying a group, but an *experiment* involves implementing a treatment with the group. This example will help students to fully understand the difference in these concepts that they have previously learned. For the purpose of graphing and organizing data, students need to discuss the difference in response and explanatory variables.

Teacher to Teacher

Connect the vocabulary to previous activities. In particular, students should see that the experiment as described involves no random sampling of people, so conclusions cannot be drawn about people in general, or even students in general. But since assignment to treatments is random, the experiment can provide evidence of causality.

3–4 Graphic Organizer, Look for a Pattern, Activating Prior Knowledge

Students will organize the data gathered from the experiment, calculate the statistics, and analyze similarities and differences in the two groups. They should notice that the "Not Fun" group is more spread out and, on average, has higher time values.

TEACHER to TEACHER

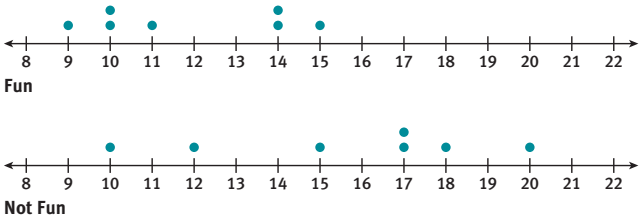
Students often forget about using the statistics that they have already learned from previous activities. Be prepared to show them how to determine one-variable statistics and to analyze their dot plots.

After completing the study for all 14 subjects, Jamie and Riley wanted to analyze the data to determine whether the subjects in the "Fun" group tended to think less time had passed than those in the "Not Fun" group. If they did, then there would be evidence that the expression, "Time flies when you are having fun!" is true.

The table below gives Jamie and Riley's data.

Group	Perceived Minutes Elapsed						
Fun	10	11	10	15	9	14	14
Not Fun	18	17	17	15	10	12	20

3. Draw dot plots of the data on the axes below to display the distributions of perceived elapsed times.



4. Compare the two distributions of estimated elapsed times in a way that addresses Jamie and Riley's research question.

Responses will focus on measures of center and spread. The "Fun" group had a mean of 11.9 minutes, a median of 11 minutes, a standard deviation of 2.41 minutes, and a range of 6 minutes. The "Not Fun" group had a mean of 15.6 minutes, a median of 17 minutes, a standard deviation of 3.51 minutes, and a range of 10 minutes. Responses will indicate that the "Fun" group had smaller measures of center and smaller spread than the "Not Fun" group.

Lesson 40-1
Random Chance

My Notes

Jamie and Riley showed their data to their classmate, Mercedes, who raised an issue they had not considered. She asked, "How do you know that your subjects would not have produced the same results regardless of which group they were in?"

5. Suppose that the perception of the passage of time was not affected by the group in which the subjects were placed. Describe how the two distributions would have appeared.
   The distributions would have looked very similar, although not exactly the same.

Mercedes suggested that Jamie and Riley consider the difference between medians of the two sets of data—median of the "Not Fun" group minus median of the "Fun" group—to compare the two groups.

6. Compute the difference in medians of the two sets. Does this indicate a difference between the two groups?
   The difference in medians is $17 - 11 = 6$ minutes. Intuitively, students may consider that such a difference in medians does imply that the treatments had an effect on the perception of the passage of time.

7. How large must the difference between the medians be to show an effect of the task assigned to each group?
   At this point, answers will be very subjective, with larger differences indicating an effect while smaller numbers would not indicate an effect.

ACTIVITY 40 Continued

5–7 Discussion Groups, Think-Pair-Share, Quickwrite In this chunk, students will recall *median* and determine if there is an effect on the perception of time based on the group a person is assigned to. Median is used for the purpose of comparison so that very large numbers and spread do not have a large effect on conclusions.

TEACHER to TEACHER

It is important for students to be able to recognize the difference in median and mean as measures of average. Ask such questions as:

- Does it matter which measure is used?
- Do the large numbers in the "Not Fun" group make a difference in the median?

CONNECT TO AP

In AP Statistics, students are expected to recognize that outliers can dramatically affect the mean of a data set, and therefore, the median should be used as a measure of center if there is a large spread in a group of data.

Developing Math Language

Be sure that students understand the meaning of *statistically significant*. It may be helpful to give an example: Michelle says that she has ESP (extra-sensory perception, or the ability to just know the answer). To determine if this is true, she takes a 10-question true/false test written in Russian (she does not know Russian). She consistently gets 8 out of 10 questions correct with repeated attempts. Does this mean that she truly has ESP? Statistically significant means that this is a very unlikely thing to happen just by chance; therefore, she truly does have ESP.

ELL Support

Discuss with students the meaning of the words *statistically and significant*, and then discuss that the statistical term means how sure you are that a difference or relationship exists.

8–9 Graphic Organizer, Look for a Pattern Students should organize the data given in the original situation in ascending order without worrying about whether they were in the "Fun" group or "Not Fun" group. This will allow a comparison of the lower time values and upper time values.

Teacher to Teacher

Students may have difficulty determining the best method to arrange the data. A discussion about spread of numbers, including the fact that ascending or descending numbers would have a significant impact on the outcome, may need to occur.

10–12 Graphic Organizer, Look for a Pattern In this chunk of items, students will organize the original numbers given in any random order. These simulations should show that there is no pattern. A student may choose to use index cards to randomly select the order.

My Notes

MATH TERMS

A statistic computed from a set of data is **statistically significant** if it would have been very unusual for the value of the statistic to be the result of chance alone.

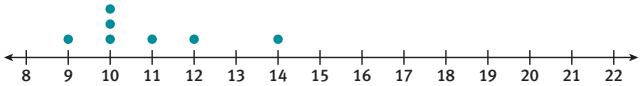
As a result of chance, two data sets cannot be expected to have exactly the same values, even if the treatments really had no effect. However, if the difference in two data sets is so great that it is very unlikely to have occurred by chance, then the results are called ***statistically significant***.

8. Using the data that Jamie and Riley collected, rearrange the "Perceived Minutes Elapsed" values in such a manner that they would be most statistically significant.

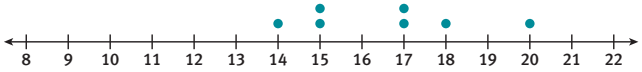
Group	Perceived Minutes Elapsed						
Fun	9	10	10	10	11	12	14
Not Fun	14	15	15	17	17	18	20

The seven smallest values should be in the "Fun" group and the largest seven values in the "Not Fun" group.

9. Create a dot plot for these groups, find the median for each, and compute the difference in medians.



Fun Median = 10 minutes



Not Fun Median = 17 minutes

Difference in medians = 17 − 10 = 7 minutes

10. Using the data that Jamie and Riley collected, rearrange the "Perceived Minutes Elapsed" values in such a manner that they would **not** be statistically significant.

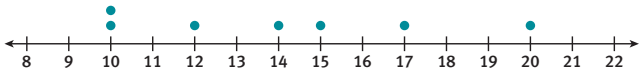
In these distributions, more diversity of answers will occur; however, there should be significant overlap of data.

Group	Perceived Minutes Elapsed						
Fun	14	10	15	10	17	12	20
Not Fun	9	15	10	17	11	18	14

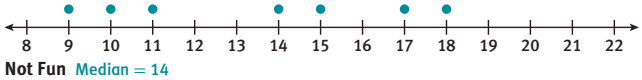
Lesson 40-1
Random Chance

My Notes

11. Create a dot plot for these groups, find the median for each, and compute the difference of the medians.



Fun Median = 14



Not Fun Median = 14

Difference in medians = 14 − 14 = 0

12. Explain why the difference in medians may be a good statistic to investigate to determine statistical significance.
In data sets that display statistically significant results, a measure of center, like the median, would be expected to be rather different. The results of Items 9 and 11 reflect this idea.

Check Your Understanding

13. Find the difference in the mean of the two test groups. Would difference in means be a good statistic for determining statistical significance in this situation?

14. Find the difference in the standard deviations of the two test groups. Would difference in standard deviations be a good statistic for determining statistical significance in this situation?

ACTIVITY 40 Continued

Check Your Understanding

Debrief students' answers to these items to ensure that they understand that a random simulation helps to determine statistical significance. By using the mean and the median as a method for comparing data, a difference that is so large that it is unlikely to have been caused just by chance can be determined. In Item 14, students may have difficulty understanding that a difference between two standard deviations shows no relevant information.

Answers

13. 15.6 − 11.9 = 3.7 minutes
If playing a video game did cause a decrease in the perception of passage of time, and if transcribing computer code did cause an increase in this perception, then the difference in means would be an appropriate statistic to investigate.

14. 3.51 − 2.41 = 1.1 minutes
The standard deviation determines the variation of the times, and the size of this statistic does not imply which group perceived the passage of time to be larger. The standard deviation may only tell which group had a more consistent experience, so the difference in standard deviations is not an appropriate statistic to answer this question of interest.

CONNECT TO AP

In AP Statistics, students learn that for a sum or difference of standard deviations to have meaning they need to use the variances, which are the squares of the standard deviations.

ASSESS

Students' answers to Lesson Practice problems will provide you with a formative assessment of their understanding of the lesson concepts and their ability to apply their learning.

See the Activity Practice for additional problems for this lesson. You may assign the problems here or use them as a culmination for the activity.

LESSON 40-1 PRACTICE

15. The treatments are different feet with which the players are taking the penalty kicks: the dominant foot and the non-dominant foot.

17. The difference in means would be better since the means are 7.45 for the dominant foot and 7 for the non-dominant foot. The median of both data sets is 7.

18. A dominant foot distribution consisting of 9, 9, 9, 9, 8, 8, 8, 8, 8, 8, 7 and a non-dominant foot distribution consisting of 7, 7, 7, 7, 7, 6, 6, 6, 5, 5, 5 would show a clear relationship between dominant and non-dominant foot successes.

19. For the difference of means 0.45 to be statistically significant, it would have to be demonstrated that such a difference was so great as to exceed a reasonable probability of occurring just by chance.

ADAPT

Check students' answers to the Lesson Practice to ensure that they understand how to compare data sets to determine statistical significance. Provide additional practice as needed.

My Notes

LESSON 40-1 PRACTICE

Carol and Alina play soccer for a local college team. Based on anecdotal evidence, they think that there is a difference in a player's success rate of taking penalty kicks with their dominant foot compared to their non-dominant foot. They would like to test this hypothesis with an experiment. Carol arranges for each of the 11 starting players on her team to take ten penalty kicks with their dominant foot and ten penalty kicks with their non-dominant foot, and records the data.

15. What are the treatments in this experiment?

Alina collected the data in the table below.

	Number of Successful Penalty Kicks										
Dominant Foot	8	7	9	9	6	7	7	8	5	9	7
Non-Dominant Foot	6	7	8	9	5	8	7	6	5	8	8

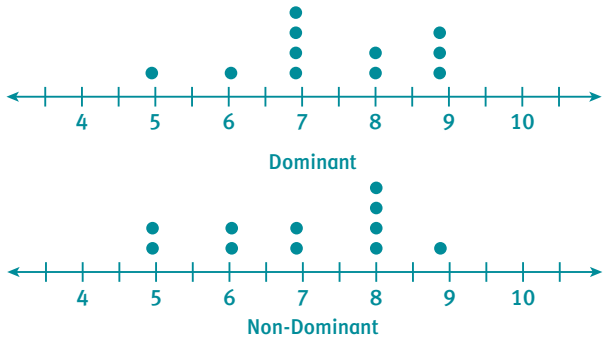
16. Draw a dot plot for each distribution. Does it seem that Carol and Alina's hypothesis is supported?

17. To test their hypothesis with this data set, which test statistic would be better: difference in medians or difference in means?

18. Using the same data values, describe two distributions that would be more supportive of the hypothesis.

19. Describe the meaning of statistical significance in this context.

16.



Answers will vary, but it seems that the dot plot for dominant foot is concentrated more to the right, while the dot plot for non-dominant foot is more uniform.

My Notes

Learning Target:

- Design and conduct a simulation to test statistical significance.

SUGGESTED LEARNING STRATEGIES: Activating Prior Knowledge, KWL Chart, Role Play, Summarizing, Paraphrasing, Think Aloud, Debriefing

Using the data values from Jamie and Riley's original study, you grouped data values in Lesson 40-1 to create data sets with median differences that were statistically significant and that were not statistically significant. However, the question remains for the original data set collected by Jamie and Riley: Is the difference of medians, 7 minutes, statistically significant in their study?

To investigate this question, create a model to randomly select data values from the original data set. This will represent a situation in which the treatment had no effect on the perception of the passage of time. Recall that the data collected from the study included responses of 9, 10, 10, 10, 11, 12, 14, 14, 15, 15, 17, 17, 18, and 20 minutes.

1. Use the *randInt* function on your calculator to choose random integers from 9 to 20. Repeat the process until you obtain seven of the data values above, without replacement. (Note that 10, 14, 15, and 17 occur multiple times, and therefore can occur the same number of times in your selection.)

a. Write those seven data values as the "Fun" values. The values that remain are the "Not Fun" values. Use the table below to organize your selections.

Group	Perceived Minutes Elapsed						
Fun	17	9	20	12	11	14	14
Not Fun	10	10	10	15	15	17	18

Sample values are listed.

b. Find the median of the "Fun" data values and the median of the "Not Fun" data values. Subtract the "Fun" median from the "Not Fun" median.

The "Not Fun" median is 15, and the "Fun" median is 14, so the difference is $15 - 14 = 1$.

TECHNOLOGY TIP

To find the *randInt* function on the TI-84, press MATH and the arrow keys to select the *PRB* menu, and select *randInt(*. The first entry is the least integer from the range you would like to sample, followed by a comma, and then the greatest integer from the range. Press ENTER and the result is an integer, chosen at random, from the range you indicated.

For example, to choose a random integer between 5 and 15, including 5 and 15, enter *randInt(5,15)*.

Lesson 40-2

PLAN

Materials:

- deck of cards that includes jokers (1 per 2 groups), or index cards with the numbers written on them to be used in place of technology for the simulation

Pacing: 1 class period

Chunking the Lesson

#1–2 #3–4

#5–6 #7–8

Check Your Understanding

Lesson Practice

TEACH

Bell-Ringer Activity

Have students solve the following problem: Juan took the SAT in September and got a math score of 570. He felt this was not a good enough score to get into his dream college. He decided to take an SAT prep class and then retook the test in December. He got a 600 on the math section. Write a sentence explaining if the SAT prep class helped Juan or not. Justify your answer.

1–2 Activating Prior Knowledge, Summarizing, Paraphrasing, Think Aloud, Use Manipulatives Students will use technology to simulate a data set using the original numbers in the study. Be sure that students know how to use the technology. Alternatively, students may instead use a deck of cards or labeled index cards.

TEACHER to TEACHER

Students often have difficulty when completing a simulation. Have students put a little mark under each number as they use it. Remind students that the numbers 13, 16, and 19 are not in the original data set, so ignore them. Also, 10, 14, and 17 can occur more than once.

CONNECT TO AP

The process of combining to get 100 sets of data is called *replication*. This is one of three important steps in determining statistical significance in AP Statistics.

3–4 Create Representations, Debriefing A quick way to gather this data would be to display the chart on the board and have students put tally marks on the chart.

My Notes

2. Repeat the process in Item 1 and record the difference of the medians in the table below.

Simulation Number	1	2	3	4	5	6	7	8	9	10
"Fun" Median – "Not Fun" Median	−1	3	3	0	−1	−3	−2	−1	−3	0

3. Combine your list of ten differences with those of your classmates so that you have at least 100 values. (The more values you have, the better your results will be). Write the results in the table below.
Answers will vary.

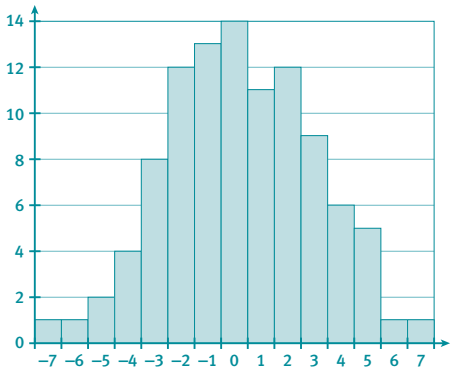
Difference of Medians	Frequency
−7	
−6	
−5	
−4	
−3	
−2	
−1	
0	
1	
2	
3	
4	
5	
6	
7	

Lesson 40-2
Testing Statistical Significance

ACTIVITY 40
continued

My Notes

4. Create a histogram with the combined class values from Item 3, and describe the shape of the distribution.



Answers will vary. A sample graph is displayed with a total of 100 data values from simulations. Distributions should be bell-shaped and approximately normal.

In Jamie and Riley's real data set, the difference between the median data values of the two groups was 7 minutes. That difference seemed rather large, but it wasn't obvious whether it was so large that it was statistically significant.

5. Describe the meaning of *statistically significant* in this context.
The difference of medians of 7 minutes would be statistically significant if it were very unlikely to occur by random chance.

6. Based on your results, what is the probability of the difference in medians being as great as 7 minutes?
This sample includes only one occurrence of 7 minutes, so the probability is $\frac{1}{100}$ or 0.01. Some students may interpret the question to include negative values ("Not Fun" has values of −6 and −7) so the probability may also be $\frac{3}{100}$. This probability is also quite small.

ACTIVITY 40 Continued

3–4 (continued) Discussion should include that the shape of the graph is approximately normal. Most students will get numbers that are close to 0 for the difference with few simulations in the area exceeding |6|.

5–6 Quickwrite, Self Revision/Peer Revision, Debriefing Students should recognize that by the time 100 or more simulations of the set of data are performed, certain values of the difference would occur more often than others. Statistically significant means that the value Jamie and Riley got would be very unlikely to occur by chance.

7–8 Quickwrite, Debriefing These items are intended to develop the idea of what should occur randomly compared to what actually occurs. If what should occur randomly is quite different from what actually occurs, then something is statistically significant. Good examples to help students to understand this concept could involve proving beyond a reasonable doubt that someone committed a crime, or whether a sports athlete who may be using enhancement drugs, fails 3 out of 12 drug tests, should therefore fined.

TEACHER to TEACHER

It is important for students to be able to draw an appropriate conclusion from data. It is also important for them to be able to explain how the data leads to that conclusion.

CONNECT TO AP

Hypothesis testing is the second of two types of inferential methods that students learn about in AP Statistics. (The third is confidence intervals).

Check Your Understanding

Debrief students' answers to these items to ensure that they understand how a simulation or random data can be used to determine statistical significance.

Answers

9. Based on the simulated distribution of the difference in sample medians under the null hypothesis, it appears that any difference of five or more would have looked unusually high, and would have been considered a statistically significant difference. (Some students may be more generous, allowing for anything 4 or higher—which is fine.)

ACADEMIC VOCABULARY

A **simulation** is a model of a real-world process in which imaginary data are generated, usually many times, to determine what results can be expected from the real-world process.

CONNECT TO AP

In AP Statistics, it is important to be able not only to draw an appropriate conclusion from data, but also to articulate a logical argument about how the conclusion follows from the data.

The ***simulation*** you just conducted is a tool to help determine what values are typical and what values are atypical, assuming that the subjects' responses would have been the same regardless of which treatment group they ended up in.

7. Based on the results of your simulation, is the 7-minute difference in medians statistically significant? Explain.
Yes. Based on the results of our sample simulation, 7 minutes seems like a very unlikely value, occurring in only one of 100 repetitions of the simulation. Therefore, the result of the experiment is unlikely due to chance alone.

8. Write a logical argument explaining what Riley and Jamie should conclude from their experiment. Your argument should draw on the assumption behind the simulation, the result of the simulation, and Jamie and Riley's actual data.
Sample answer: Suppose that each of the 14 subjects in Jamie and Riley's experiment was going to give the same estimate of elapsed time, regardless of which treatment group he or she was put in. In this case, a difference between medians of 7 minutes would be very unusual; a difference that big occurred only one time out of 100 simulations of the experiment. While one possible explanation for the difference is that it was just a very unusual occurrence, a more reasonable explanation is that the treatments did have an effect on the subjects. Performing the task in the "Fun" group really did have the effect of subjects perceiving less elapsed time than those in the "Not Fun" group. Jamie and Riley have good evidence that the adage is true: "Time flies when you are having fun!"

Check Your Understanding

9. In Jamie and Riley's experiment, are there other possible values that their test statistic might have taken which would have led them to the same conclusion—that "time flies when you are having fun"?

10. Write a conclusion that Jamie and Riley could have drawn if the difference between the medians in their study had been 3 minutes instead of 7 minutes.

10. Sample response: In the data, the subjects in the "Fun" group gave a median estimated elapsed time that was three minutes shorter than the median estimated elapsed time given by subjects in the "Not Fun" group. Although this may initially suggest that playing the video game had the effect of subjects perceiving smaller time estimates than they would have if they had transcribed code, we can see from a simulation that in fact a difference of three minutes or more is not all that unusual, occurring 22% of the time, assuming that the treatments do not have any effect. Therefore, the difference of three minutes is not statistically significant, and our experiment did not provide evidence that "Time flies when you are having fun!"

My Notes

LESSON 40-2 PRACTICE

Use the following for Items 11–13.

In the experiment described in this activity, Jamie and Riley chose to use the difference between the median responses of subjects in the two groups as their test statistic. Suppose instead that they had decided to look at the ratio of the means from the two groups by dividing the mean perceived time in the "Not Fun" group by the mean perceived time in the "Fun" group.

11. Compute the mean perceived times for the "Not Fun" group and the "Fun" group, and then write the ratio.

12. Interpret this ratio in terms of the context of time perception between the "Not Fun" and "Fun" groups.

13. After completing many simulations, what would Jamie and Riley do next to test their hypothesis?

Use the following for Items 14–16.

Recall that in Jamie and Riley's study, they decided to interrupt the subjects' activity at 13.5 minutes. Suppose they had decided instead to interrupt them after 17 minutes.

14. What would have been different about the data?

15. How would the test statistics of difference in medians and ratio of means have changed?

16. Would Jamie and Riley's conclusions be different if the actual time that participants were involved with their activity were increased to 17 minutes? Explain your reasoning.

15. There would likely be little change in the test statistics as the medians and the means in both groups may have increased by similar amounts. (An additional three minutes and thirty seconds is unlikely to have changed the perception of the passage of time significantly. To know for certain, however, another study would have to be conducted to determine if perception of the passage of time is affected by the actual duration of time.)

16. Jamie and Riley would likely have come to the same conclusions. They compared the two groups to one another, not to the number 13.5 itself. Their question of interest was not how accurately people estimated the passage of time, but whether one's activity had an influence on one's estimate of the passage of time.

ASSESS

Students' answers to Lesson Practice problems will provide you with a formative assessment of their understanding of the lesson concepts and their ability to apply their learning.

See the Activity Practice for additional problems for this lesson. You may assign the problems here or use them as a culmination for the activity.

LESSON 40-2 PRACTICE

11. They would have computed, for their actual data, the ratio of the means for the two groups, which turns out to be $\frac{15.57}{11.85} = 1.31$.

12. For every one minute of perceived time passage for the "Fun" group, the "Not Fun" group perceived 1.31 minutes of time passage.

13. Jamie and Riley would have looked at the distribution of ratios of the mean to see whether their observed test statistic of 1.31 was unusually high or too high to be plausibly attributable to chance alone.

14. The actual data value would have been larger as people would have perceived that a longer period of time had passed.

ADAPT

Check students' answers to the Lesson Practice to ensure that they understand how to analyze data and use simulations to make decisions in context about statistical significance. Interested teachers, or teachers with interested students, may want to ask students to actually conduct a randomized experiment on their own. An experiment might involve something as simple as testing whether students are as good at texting while standing on one foot as opposed to two feet, or whether they can remember more words in a list when they've seen them but not heard them compared with hearing them but not seeing them. Letting students take ownership of an experiment (with guidance) can be an effective way to engage students who are otherwise less engaged with the course content. The conclusion of such an experiment could simply be a one-page written description of the research question, how the experiment was executed, and the conclusions.

ACTIVITY PRACTICE

CONNECT TO AP

In AP Statistics, students are required to reject or fail to reject a null hypothesis. Statistical significance is the basis of this concept.

1. The question of interest is whether the success rate of throwing the toy ball into a basket is influenced by the knowledge of difficulty of the task.
2. The treatments are telling subjects that "One-fourth of all people have made it" and "One-fourth of all people have missed it."
3. A positive test statistic would indicate that those who thought the task was easy were more successful than the other group. A negative test statistic would indicate that those who thought the task was difficult were more successful than the other group. A test statistic of zero would indicate that there was no difference in the success rates of both groups.
4. $\frac{29}{48} \approx 0.604$ and $\frac{25}{52} \approx 0.481$, so the difference of the two proportions is 0.123.
5. They would be able to conclude that being told that a shot is easy ("most people made it") makes a person more likely to make the shot than if they'd been told that the shot is hard ("most people miss it").
6. They would only be able to conclude that their data did not provide any evidence that telling a person that a shot is easy or hard has any influence over whether or not they will make the shot. It is *not* correct to conclude that "telling a person that a shot is easy or hard has no effect on whether they'll make the shot". Absence of evidence is not the same thing as evidence of absence.
9. Sample response: The count of 15 babies reaching for the nice puppet is somewhat on the high side in the first distribution, but not so unusually high that it would be very surprising for it to occur by chance alone. Perhaps babies may reach for the nice puppet more often than chance alone would dictate, but the given data do not provide statistically significant evidence of that.

ADDITIONAL PRACTICE

If students need more practice on the concepts in this activity, see the Teacher Resources at SpringBoard Digital for additional practice problems.

ACTIVITY 40 PRACTICE

Write your answers on notebook paper.
Show your work.

Lesson 40-1

Abraham and Luis are interested in conducting an experiment to see whether people's ability to successfully toss a toy ball into a basket is influenced by their belief that others found the task difficult or easy. They position themselves in a central location at their school, place a basket 15 feet away from a spot marked "X," and ask volunteers to try to make the basket. They randomly choose subjects to participate and randomly tell them one of two statements: "So far, only one-fourth of people have made it" or "So far, only one-fourth of people have missed it." They repeat this for a total of 100 subjects. Their data are summarized below.

	Number Who Made Shot	Number Who Missed Shot	Total
Told That Most People Made It	29	19	48
Told That Most People Missed It	25	27	52

1. What is the question of interest in this study?
2. What are the treatments imposed by Abraham and Luis in this experiment?
3. Abraham and Luis decided that their test statistic is the difference between the proportion of people who made the shot in the group that was told the task was easy and the proportion of the people who made the shot in the group that was told the task was difficult. Interpret the meaning of a positive test statistic, a negative test statistic, and a test statistic of zero in this context.
4. Compute the test statistic for the results of the experiment.
5. If it was determined that the test statistic was statistically significant, what would Abraham and Luis be able to conclude?
6. If it was determined that the test statistic was not statistically significant, what would Abraham and Luis be able to conclude?

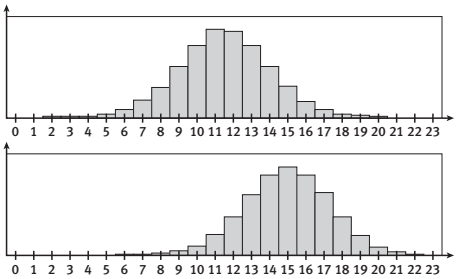
Lesson 40-2

7. Abraham and Luis's teacher provided them with 200 beads, 100 red and 100 white, of which the only difference was their color. Describe how Abraham and Luis could use those beads to create a simulation to determine whether their test statistic is statistically significant. Be sure to identify what the beads represent.

Use the following for Items 8 and 9.

In a study designed to determine whether babies have an innate sense of morality, babies were shown two puppet shows in a random order: one of them had a puppet being nice, and the other had a different puppet being mean. The babies were then given the opportunity to reach for either the nice puppet or the mean puppet, and the researchers recorded which puppet the babies reached for. Suppose that out of 23 babies in the study, 15 of them reached for the nice puppet.

8. One of the distributions below shows the probability distribution of the number of babies who would reach for the nice puppet if, in fact, babies had no sense of morality and were reaching for a puppet at random. Which distribution is it, and how do you know?



MATHEMATICAL PRACTICES

Construct Viable Arguments and Critique the Reasoning of Others

9. Using the distribution you picked in Item 8 and the observed 15 out of 23 babies reaching for the nice puppet, what conclusion should be drawn, and why?

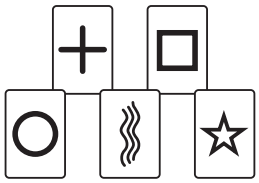
7. Answers will vary, but should be similar to the simulation that follows.
Let red beads represent people who make their shot, and white beads represent people who miss their shot. Since $29 + 25 = 54$ people made the shot and $19 + 27 = 46$ people missed the shot, then make piles of 54 red beads and 46 white beads. Mix the piles up and then, without looking at their colors, count out 48 beads to represent people who were told the shot was easy and 52 beads to represent people who were told the shot was hard. In each group, compute the proportion of the beads that are red. Then compute the difference. Record that difference, mix up the beads, and repeat the whole process many times. Make a histogram or dot plot of the distribution of the simulated differences. If the number 0.123 appears to be unusually high in that distribution, then the difference is statistically significant; if not, then it is not statistically significant.
8. The first distribution shows how many babies would pick the nice puppet if they were picking at random. That represents an assumption that babies are just as likely to pick the nice puppet as the mean one, so you would expect about half of the 23 babies—11 or 12 of them—to pick the nice puppet by chance. And you would expect counts larger than that to be as likely as counts lower than that. The first data display is symmetric and centered on about 11 or 12, so it fits that description. The other data display has more than half of the babies choosing the nice puppet with a much greater probability than ought to be the case if the babies are picking at random.

Simulations, Margin of Error, and Hypothesis Testing

Embedded Assessment 2
Use after Activity 40

PSYCHIC OR JUST HOT AIR?

1. "Zener cards" are used to test whether someone has extrasensory perception (ESP). Each card has one of five distinct images on it:



Suppose that a subject is presented with a random assortment of 12 such cards and is asked to identify the images without looking at them. He correctly identifies 6 out of the 12 cards.

a. Given a random card, what is the probability of correctly identifying the image on the card?

b. Use the random digits table below, with 0 and 1 representing correct identifications and digits 2-9 representing incorrect identifications, to perform ten different simulations. Beginning with row 113, record the number of successes for each trial, and make a dot plot of your results.

111	81486	69487	60513	09297	00412	71238	27649	39950
112	59636	88804	04634	71197	19352	73089	84898	45785
113	62568	70206	40325	03699	71080	22553	11486	11776
114	45149	32992	75730	66280	03819	56202	02938	70915
115	61041	77684	94322	24709	73698	14526	31893	32592

c. What conclusion does this data support?

2. An engineer developed a treatment that he hoped would make the fabric of a hot-air balloon last longer. Out of 9 volunteer balloonists with new balloons, he randomly selected 4 to get no special treatment, and 5 to get their fabric treated. The table below shows how many balloon-hours the nine balloons lasted.

No Special Treatment	520	610	435	443	
Received the Special Treatment	496	639	550	622	600

a. Find the difference in the means of the two groups.

b. Each balloon-hour total is written on an index card. Describe a simulation using these cards that could help determine the statistical significance of the difference of the means you found in part a.

c. Describe a manner in which the results of the simulation would allow you to reasonably conclude that the difference in the means was not statistically significant.

3. A regular survey asks a random sample of 1070 American adults whether they approve of the job the President of the United States is doing. The margin of error in the proportion of people who say "yes" is stated to be ±3 percentage points.

a. Suppose such a survey yielded a proportion of 0.45. Explain what that means in everyday language.

b. How could the survey be conducted differently to reduce the margin of error?

Common Core State Standards for Embedded Assessment 2

HSS-IC.A.1	Understand statistics as a process for making inferences about population parameters based on a random sample from that population.
HSS-IC.A.2	Decide if a specified model is consistent with results from a given data-generating process, e.g., using simulation.
HSS-IC.B.4	Use data from a sample survey to estimate a population mean or proportion; develop a margin of error through the use of simulation models for random sampling.
HSS-IC.B.5	Use data from a randomized experiment to compare two treatments; use simulations to decide if differences between parameters are significant.
HSS-IC.B.6	Evaluate reports based on data.

Embedded Assessment 2

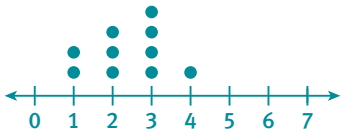
Assessment Focus

- Simulation of random processes
- Testing the truth of a conjecture
- Statistical significance
- Margin of error

Answer Key

1. a. The probability of correctly identifying the image on the card is 0.2, or one out of five.

b. The results of the simulation are 2, 2, 3, 4, 1, 2, 3, 3, 3, 1. The dot plot is shown.



c. Since no data points are close to six, it seems very unlikely that six successes in twelve trials with a probability of 0.2 is due to chance. Therefore, it is safe to reject the hypothesis that the subject does not have ESP.

2. a. The difference in means, 581.4 for treated balloons and 502 for untreated balloons, is 79.4 balloon-hours.

b. The cards would be shuffled, and five cards would be chosen at random to represent the balloons with treated fabric. The remaining cards would represent balloons with untreated fabric. The mean of each group would be determined, and the difference in the means would be computed. The process would be repeated several times, and a distribution of the differences of the means would be created.

c. If the results of the simulation indicated that the difference of 79.4 balloon-hours was not very unusual and could have occurred by chance, then the difference in the means is not statistically significant.

3. a. In a simple random sample of 1070 American adults, the actual population proportion will likely be within 0.03 of the sample proportion. The actual population proportion for this sample will very likely be between 0.42 and 0.48.

b. The number of subjects in the sample could be increased; as the number of subjects in the sample increases, the margin of error decreases.

Embedded Assessment 2

Teacher to Teacher

You may wish to read through the scoring guide with students and discuss the differences in the expectation at each level. Check that students understand the terms used.

Embedded Assessment 2
Use after Activity 40

Simulations, Margin of Error, and Hypothesis Testing

PSYCHIC OR JUST HOT AIR?

Scoring Guide	Exemplary	Proficient	Emerging	Incomplete
	The solution demonstrates these characteristics:			
Mathematics Knowledge and Thinking (Items 1, 2, 3)	• Clear and accurate understanding of significance testing using a table of random digits or a simulation • Clear and accurate understanding of margin of error and survey design principles	• A functional understanding and accurate interpretation of significance testing using a table of random digits or a simulation • A functional and mostly accurate understanding of margin of error and survey design principles	• Partial understanding and partially accurate interpretation of significance testing using a table of random digits or a simulation • Partial understanding and partially accurate work with margin of error and survey design principles	• Little or no understanding and inaccurate interpretation of significance testing using a table of random digits or a simulation • Little or no understanding and inaccurate work with margin of error and survey design principles
Problem Solving (Items 1, 2, 3)	• An appropriate and efficient strategy that results in a correct answer	• A strategy that may include unnecessary steps but results in a correct answer	• A strategy that results in some incorrect answers	• No clear strategy when solving problems
Mathematical Modeling / Representations (Items 1, 2)	• Clear and accurate understanding of how to apply simulations and random digit tables to analyze real-world scenarios	• Mostly accurate understanding of how to apply simulations and random digit tables to analyze real-world scenarios	• Partial understanding of how to apply simulations and random digit tables to analyze real-world scenarios	• Inaccurate or incomplete understanding of how to apply simulations and random digit tables to analyze real-world scenarios
Reasoning and Communication (Items 1, 2, 3)	• Precise use of appropriate math terms and language to describe margin of error and how to reduce it in a survey • Clear and accurate explanation of methods to determine statistical significance	• Adequate description of margin of error and how to reduce it in a survey • Adequate explanation of methods to determine statistical significance	• Misleading or confusing description of margin of error and how to reduce it in a survey • Misleading or confusing explanation of methods to determine statistical significance	• Incomplete or inaccurate description of margin of error and how to reduce it in a survey • Incomplete or inadequate explanation of methods to determine statistical significance