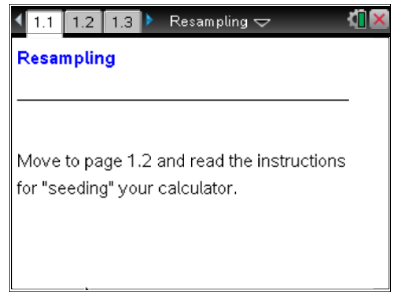


Math Objectives

- Students will recognize that a sample can be used to generate a simulated sampling distribution from which a hypothesis test can be completed.
- Students will recognize that in some situations it is possible to carry out a hypothesis test without any assumptions about the distribution of the population.
- Students will use appropriate tools strategically (CCSS Mathematical Practices).
- Students will reason abstractly and quantitatively (CCSS Mathematical Practices).

Vocabulary

- hypothesis test
- mean
- median
- normal distribution
- parameter
- population
- sample
- sampling distribution
- statistic

About the Lesson

- This lesson involves approximate sampling distributions obtained from simulations based directly on a single sample. The focus of the lesson is on conducting hypothesis tests in situations for which the conditions of more traditional methods are not met.
- As a result, students will:
  - Develop sampling distributions of sample statistics from samples in a two-treatment scenario and complete associated hypothesis tests.

TI-Nspire™ Navigator™ System

- Transfer a File
- Use Screen Capture to compare different graphical displays.

Prerequisite Knowledge

- familiarity with sampling distributions
- understanding of the logic of hypothesis tests

Related Activities

- Sampling Distributions
- Why t?

TI-Nspire™ Technology Skills:

- Download a TI-Nspire document
- Open a document
- Move between pages
- Grab and drag a point

Tech Tips:

- Make sure the font size on your TI-Nspire handhelds is set to Medium.
- You can hide the function entry line by pressing ctrl G.

Lesson Files:
Student Activity
Resampling_Student.pdf
Resampling_Student.doc
TI-Nspire document
Resampling.tns

Visit www.mathnspired.com for lesson updates and tech tip videos.

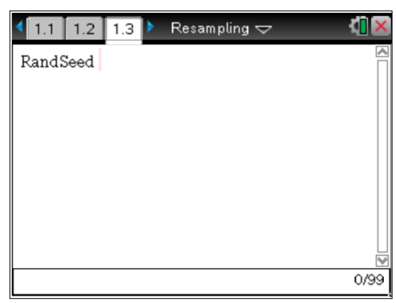
Discussion Points and Possible Answers

Teacher Tip: This lesson develops one resampling technique, an approximation to the permutation test, for hypothesis testing without the customary assumptions needed for the two-sample t-test. Another resampling method, the bootstrap, is not discussed.

Tech Tip: Page 1.2 gives instructions for seeding the random number generator on the handheld. Page 1.3 is a Calculator page for the seeding process. Ensuring that students carry out this step will prevent students from generating identical data. (Syntax: RandSeed #, where # is a number unique to each student.).

Teacher Tip: Once students have seeded their random number generators, they do not have to do it again unless they have cleared all of the memory. But it is important that this be done if the memory has been cleared or with a new device, as otherwise the "random" numbers will all be the same as those on other similarly cleared devices.

Move to pages 1.2 and 1.3.



A class of 18 students decided to test the relative effectiveness of two brands of mosquito repellant. In particular, their research question was "Is Brand B more effective than Brand A in preventing mosquito bites?" Note that this creates a hypothesis test in which the alternative hypothesis is one-sided, favoring Brand B. Each member of the class tossed a coin to decide which brand to test, applied the appropriate repellant according to the manufacturer's instructions, then joined the rest of the class at the sports stadium for the evening's athletic event. Times until the first mosquito bite were recorded (in minutes).

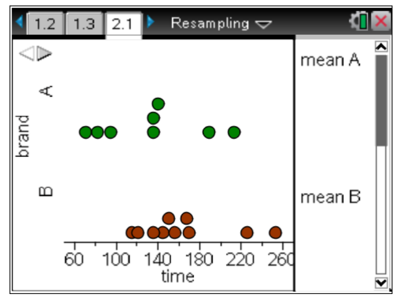
1. Do you think that all the students who test brand A will get exactly the same protection times? If "yes," explain why. If "no," identify possible sources of variability.

   Sample Answer: No. There is always random variation. Since different students are the subjects, any variation among students (body temperature, physical activity, recent food consumption, etc.) will contribute to the variation. In addition, the amount of repellant used is likely to differ somewhat from person to person.

2. If the two brands protect, on average, equally effectively, do you think that the student sample data will show identical mean times before the first bite? If “yes,” explain why. If “no,” tell what could cause the difference.

   Sample Answer: No. There is always random variation. As noted in question 1, students vary. It is possible that some students are more attractive to mosquitoes and others are less attractive. So if the coin toss results in the mosquito-attractive and mosquito-unattractive students not being equally split between the two brands there could be an apparent difference in the data even when the repellants are essentially identical.

Move to page 2.1.



3. Page 2.1 displays the data obtained from the students. Based only on the comparative dotplots, do you believe that the average protection times for the two brands really differ, or could these results be the result of chance variation? Explain your reasoning.

   Sample Answer: It’s hard to say. The plot for the sample from brand A seems shifted a little to the left of that for brand B, but there is a lot of overlap between the two brands. In addition, there is a lot of variability within each brand, so maybe there is really no difference.

4. a. The table below provides space to write the values of the time in minutes before the first bite for each of the 18 students. Hover over each point in the dotplots on Page 2.1 to obtain the numerical values of all the data. Record each value in the top row below, and record whether it was brand A or brand B in the second row.

   Sample Answer: Protection time in minutes

70	82	95	135	135	140	189	213	115	121	135	145	150	156	167	170	226	253
A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B

   b. Click the arrow on Page 2.1 to display the mean protection times of the two brands as vertical lines on the plot. Click on each line to see its numerical value. Record these values.

   Answer: mean A =132.375, mean B = 163.8

   c. Use the dotplot or your table of values from part a to determine the median protection time of each brand. Record your results.

   Answer: median A = 135, median B =153.

Teacher Tip: The fundamental assumption is that the two groups, here A and B, are really from a single population and that the brand labels represent one (observed) of the many possible brand assignments that could occur by chance. If the two populations really are essentially alike (which is our null hypothesis), then the only mechanism leading to any particular protection time's being associated with a particular brand is the coin toss. That is, those 18 times in the table for Question 4 would still have been the data, but they could have been "rearranged" as far as A and B are concerned. Discuss student answers to 5b to ensure that this idea is clear. It is essential that students connect the null hypothesis of "no difference" with the random reassignment that will be used to generate a sampling distribution.

Every hypothesis test compares data from an observed sample or treatment allocation to a sampling distribution of what would happen if the null hypothesis were true.
Suppose the null hypothesis is 'no essential difference in the effectiveness of the two repellants, at least as far as protection time is concerned.' We need to obtain a sampling distribution of what can happen under these conditions.

5. a. If the null hypothesis is true, explain why it would be reasonable to think of the two sets of protection times as really being just one data set.

   Sample Answer: Well, if the two repellants have no essential differences, then it's like we really have only one repellant. That means that the numbers would have been the same whether or not there were two repellants, so it's like we sampled from one population, not two.

   b. If the null hypothesis is true, what would explain any apparent differences between the two plots in the students' sample data?

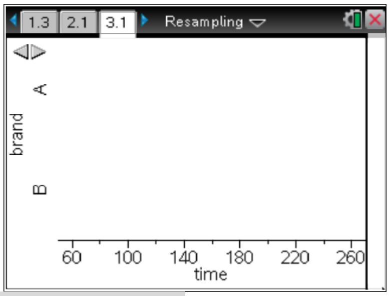
   Sample Answer: Luck. The assignment of brands to students was random, so the difference that is visible might be due to nothing but random variation or the unlucky placement of students who are attractive to mosquitos.

If the two brands are essentially alike (which is our null hypothesis), then the only mechanism leading to any particular protection time's being associated with a particular brand is the coin toss that the class did to decide which repellant to use. That is, those 18 times would still have been the data, but they could have been "reassigned" as far as A and B are concerned, had the coin tosses come out differently. So, looking back at your table in 4a, you could use coin tosses to assign brands A and B to those same numbers, and you would have another possible outcome of the class's experiment.

From the point of view of a hypothesis test, the question now is whether the observed data the 18 students did get is “typical” or “unusual” when compared to all the other sample data they could have obtained if the null hypothesis were true.

> **Teacher Tip:** The definition of “typical” or “unusual” can be left undefined, to be determined by each student separately. However, you might also want to have a discussion within the class after several simulations have been done, say after Question 7, and come to some general agreement regarding what you will call “typical.” Such a collective decision will make “reject the null” and “fail to reject the null” conclusions more uniform within the class.

Move to page 3.1.



The arrow on the left side of Page 3.1 performs the equivalent of 18 coin tosses and assigns the corresponding As and Bs to the same numbers you listed in 4a. Thus, each new click gives another possible set of data.

> **Teacher Tip:** Question 4 provides a visual representation of the “observed” permutation. If time permits, it is worthwhile to have students toss coins to write another possible permutation underneath their answer to Question 4 to complete the visualization of exactly what the permutation test is doing before turning that work over to the TI-Nspire in Question 6.

> **Teacher Tip:** In Question 6, students should realize that the majority of their plots will be reasonably consistent with the null hypothesis. That’s the point—the null hypothesis was used as the basis for the simulation. However, sampling variability is still present, so it is possible (likely when considering your entire class) to get samples that do seem to lead us to reject the null hypothesis. (This could lead to a discussion of Type I error, which is not part of this lesson.)

TI-Nspire Navigator Opportunity: ***Screen Capture***
See Note 1 at the end of this lesson.

6. a. Click the arrow once, and then click on each vertical line to see the mean of each of the two sets. Record the difference between means, $\bar{x}_B - \bar{x}_A$. Do the new comparative dotplots indicate that the average protection times for the two brands differ? Explain your reasoning.

   Sample Answer: Not really. The two dotplots overlap a lot and seem to have very similar centers, shapes, and spreads.

   b. Click the arrow a few more times. Record the difference of means, $\bar{x}_B - \bar{x}_A$, for each new rearrangement. Decide whether each new pair of comparative dotplots indicates that the repellants' average protection times differ by brand. Explain your reasoning.

   Sample Answer: None seem to indicate any difference. The two dotplots overlap a lot and seem to have very similar centers, shapes, and spreads.

   c. Think about how the new reassignments are being made. How is the major assumption behind these reassignments related to the null hypothesis?

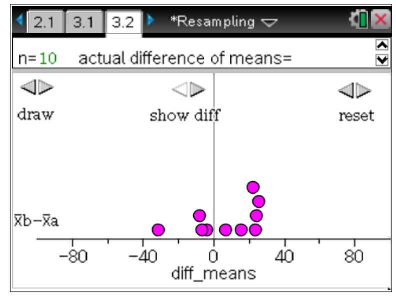
   Sample Answer: They say the same thing. The assumption for creating new reassignments is that As and Bs are assigned randomly, and the null hypothesis says that there is no systematic difference between the A numbers and B numbers.

The logic of hypothesis testing requires that we compare some statistic from the observed data to the sampling distribution of that same statistic obtained when the null hypothesis is known to be true. If the observed statistic differs from what is typical in the sampling distribution for a true null hypothesis, then either the sample was unlucky (unusual just by chance) or the null hypothesis is not actually true.

> **Teacher Tip:** In the traditional two-sample t-test, the sampling distribution is for a t-statistic calculated from the sample means and standard deviations, assuming a normally-distributed sampling distribution, and assuming the null hypothesis is true. The resampling approach replaces these assumptions with the null hypothesis and the simulated sampling distribution.

Let the test statistic be the difference of sample means, $\bar{x}_B - \bar{x}_A$, but make no assumptions about the shape of the population distribution. If the two populations are essentially alike (the null hypothesis), then the simulation you carried out for Question 6 produces exactly the samples that lead to the necessary sampling distribution.

Move to page 3.2.



Page 3.2 displays a dotplot of the differences $\bar{x}_B - \bar{x}_A$ for the samples you examined on Page 3.1. Hover over the points on the plot to verify that they agree with what you recorded in Question 6.

7. a. Click the "draw" arrow at the top left of Page 3.2 to continue generating values in the simulated sampling distribution of $\bar{x}_B - \bar{x}_A$ until you have 100 simulated samples. (Note the counter in the top panel.) Remember, this simulation uses a process that guarantees the null hypothesis is true—the only difference between As and Bs is random assignment. Describe the simulated sampling distribution of differences of means, including noting the range of values that seem "typical."

Sample Answer: The simulated sampling distribution is roughly symmetric and mound-shaped, centered at 0, with typical values between -40 and 40.

b. Click the "show diff" arrow at the top center of Page 3.2 to display the difference of means from the actual class data, both as a vertical line in the plot and as a numerical value above the plot. Does the value from the class seem unusual when compared to the possible values that make up the sampling distribution for a true null hypothesis? Explain.

Sample Answer: The actual difference of means is 31.425. It seems to fall within the "typical" values of the sampling distribution, so I don't think it is particularly unusual. Since 12 of my simulated differences are larger than 31.425, the approximate *P*-value of the class observation is 0.12, which also indicates that their result is not unusual.

> **Teacher Tip:** The simulated *P*-value may be obtained as the proportion of simulated differences that are "at least as extreme as" the observed difference. Since this study used a one-sided alternative (refer to the initial research question), "at least as extreme as" means "greater than or equal to."

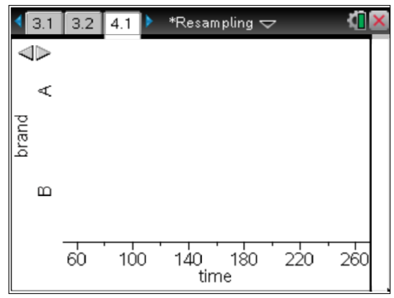
c. Based on your comparison of the class data to the simulated sampling distribution, state your conclusion for the test of our null hypothesis.

Sample Answer: I fail to reject the null hypothesis and conclude that it is possible that the observed difference of 31.425 could reasonably have been due only to chance.

8. One advantage of this method of completing a hypothesis test is that it can be applied to statistics other than means. For example, you observed the medians of the two data sets in 4c. Describe how you could conduct a hypothesis test of these data using medians.

Sample Answer: I could build a simulated sampling distribution of sample medians using the same method of constructing samples to include—random assignment of As and Bs to the original numbers. Then compare the actual difference of medians from the real sample to the sampling distribution of possible differences of medians to see whether it is unusual or not.

Move to page 4.1.

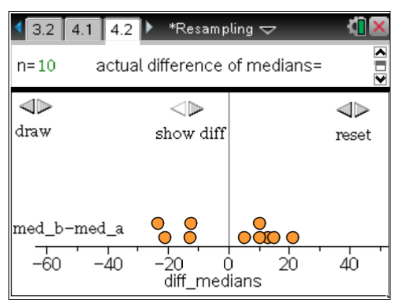


9. Click the arrow a few times. Record the differences of medians, $\tilde{x}_B - \tilde{x}_A$, for each new rearrangement. Decide whether each new pair of comparative dotplots indicates that the populations' average times before the first bite for the two brands differ. Explain your reasoning.

Sample Answer: None seem to indicate any difference. The two dotplots overlap a lot and seem to have very similar centers, shapes, and spreads.

Teacher Tip: For Question 10, student answers will vary since each simulation generates different data. It would be good to remind students that they are using an approximate sampling distribution, not the actual sampling distribution. It would be impossible to compute and plot "all possible rearrangements" for the actual sampling distribution.

Move to page 4.2.



10. a. Click the arrow at the top left of Page 4.2 to continue generating values in the simulated sampling distribution of $\tilde{x}_B - \tilde{x}_A$ until you have 100 simulated samples. Describe the simulated sampling distribution of differences of medians.

Sample Answer: The simulated sampling distribution is slightly skewed right and mound-shaped, centered at 0, with typical values between -20 and 20. Note: Since results are based on simulations, descriptions will vary from student to student.

b. Click the arrow at the top center of Page 4.2 to display (both as a vertical line and a numerical value) the difference of medians from the actual class data. Does the value from the class seem unusual when compared to the possible values that make up the sampling distribution for a true null hypothesis? Explain.

Sample Answer: The actual difference of medians is 18. It seems to fall among the “typical” values of the sampling distribution, so I don’t think it is particularly unusual. The approximate P-value, obtained by counting points beyond 18, is 0.16, which also indicates that the class data are not unusual.

c. Based on your comparison of the class’s observed sample median to the simulated sampling distribution, state your conclusion for the test of the null hypothesis.

Sample Answer: I fail to reject the null and conclude that it is possible that the observed difference between the two treatment groups was due only to chance.

TI-Nspire Navigator Opportunity: *Quick Poll*
See Note 2 at the end of this lesson.

Wrap Up

Upon completion of the lesson, the teacher should ensure that students are able to understand:

- A sample can be used to generate a simulated sampling distribution from which a hypothesis test can be completed.
- A hypothesis test can be performed in some situations without any assumptions about the population, such as normality.

Assessment

Answer each of the following:

1. What is an advantage of doing a hypothesis test using resampling?

Sample Answer: Resampling does not rely on assumptions about the sample’s coming from a normally distributed population so you can use a it without knowing anything about the population.

2. What null hypothesis is used in a permutation test regarding two populations?

   Sample Answer: The hypothesis is that the data come from two identical populations. This is equivalent to assuming that there is only a single population, so you are analyzing the behavior of two “groups” drawn from the same population.

3. Caro announced it would even be possible to have one of the elements in the simulated sampling distribution for all of the Brand A people (in the observed data) to be assigned Brand B and vice versa. What would you say to Caro?

   Sample Answer: Caro is right; that is one of the possible arrangements that could occur from the coin tosses that form the allocation of brands to people.

4. Fred stated, "In a permutation test, you are just randomly assigning half of the data to one group and half to the other and simulating a distribution of the differences for some statistic." Do you agree with Fred? Why or why not?

   Sample Answer: Not necessarily. Assignment of each data value to one group or the other is determined randomly, but that might not result in half of the values in one group and half in the other.

TI-Nspire Navigator

Note 1

Name of Feature: Screen Capture

A Screen Capture can be used to compare samples based on the random reassignment of values and labels.

Note 2

Name of Feature: Quick Poll

Quick Poll can be used to check for student understanding throughout the activity. The suggested assessment questions could be used to check for understanding at the end of the activity.