Math Objectives

- Students will recognize that the alpha value (significance level of a test) is the relative frequency for sample statistic values that lead to a "reject the null" conclusion, given that the null hypothesis is actually true.
- Students will recognize that a given sample mean can lead to a "reject the null" or "fail to reject the null" depending on the alpha level.
- Students will reason abstractly (CCSS Mathematical Practices).

Vocabulary

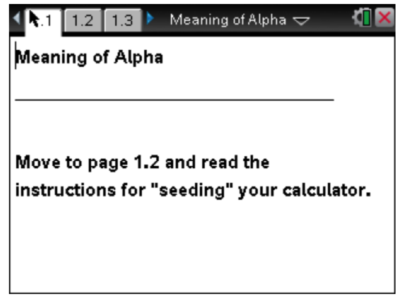
- alpha level
- population
- samples
- sampling distributions of sample means
- significance level

About the Lesson

- This lesson involves beginning with a null hypothesis specifying the mean of a normally distributed population with a given standard deviation.
- As a result, students will:
  - Generate a sample and sample mean and predict whether the sample might reasonably have come from the hypothsized population, considering the criteria they would use to make a decision.
  - Consider a fixed alpha level (0.1) and predict the likelihood of getting by chance an outcome at least as extreme as the value determining the border of the rejection region if the null hypothesis is true.
  - Generate 100 samples, and observe how many fall into the rejection region to observe that the decision to reject or not to reject the null hypothesis will change depending on alpha.

TI-Nspire™ Navigator™ System

- Use Screen Capture to compare student results.
- Send the .tns file to students.
- Use Quick Poll to determine student understanding.



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.
- In the *Graphs & Geometry* application, you can hide the function entry line by pressing ctrl G.

Lesson Materials:
Student Activity

- Meaning_of_Alpha.pdf
- Meaning_of_Alpha.doc

TI-Nspire document

- Meaning_of_Alpha.tns

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

Prerequisites

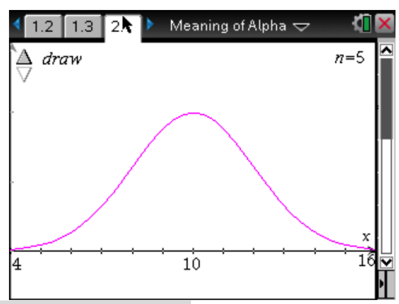
Students should be familiar with the concepts of sampling and sampling distributions. They should understand the relationship between probability distributions and area. This activity does not assume students have familiarity with *p*-values; it can be used either before or after the activity *What is a p-value?* with some slight modifications.

Discussion Points and Possible Answers

Tech Tip: Page 1.2 gives instructions on how to seed the random number generator on the TI-Nspire. 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 or else the "random" numbers will all be the same as those on other similarly cleared devices.

Consider the following hypothesis test:
H_0: $\mu = 10$
H_a: $\mu > 10$
The graph represents a normal population with a standard deviation of 2 that satisfies the null hypothesis.



Teacher Tip: The graph on Page 2.1 is the normal distribution with mean 10 and standard deviation 2. This graph rescaled is the graph from the top work area on Page 2.2. Note that it is important for students to use the same scale when they are comparing different distributions from the normal curve family. Students should recall from the *Sampling Distribution* activity or from other prior work that the sampling distribution of sample means is normally distributed with mean 10 and standard deviation $\frac{\sigma}{\sqrt{n}}$ *or* $\frac{2}{\sqrt{5}}$, which is represented in the graph from the bottom work area on Page 2.2.

> **Teacher Tip:** The lesson assumes H_a: $\mu > 10$ but could be adapted for H_a: $\mu < 10$.

1. Suppose you were to draw a sample of size 5 from the population described by the null hypothesis. Sketch what you would expect the sample to look like.

Sample Answers: Students will have a variety of sketches, but most will probably center around 10.

> **Teacher Tip:** The objective here is to get students to begin thinking about what it means to have a sample that is typical for a hypothesized population.

Move to page 2.1.

2. Draw a sample of size 5 by clicking the arrow (▲).
   a. How do the points you generated compare to your prediction in Question 1?

      Sample Answers: Some students will have a set of points that is similar to their sketch, but others might have sets of points that have much greater or smaller ranges or are centered at one or the other end of the distribution rather than at 10.

   b. Describe the distribution of those points.

      Sample Answers: Student answers should include a reference to the center and spread of their own distribution. For example, the set of points went from about 7.5 to 10, and the median was at 9.

   c. How does your distribution compare with your classmates' distributions?

      Sample Answers: Some students might have very similar distributions, but others might have very different distributions. Students should use center and spread in their comparisons: "Ours were not that different except his was a bit more to the right; I had a mean of 9.8 and a range of 2.6 (from 8.6 to 10.2), and he had a mean of 10.2 and a range of 3.6 (from 8.4 to 11)."

3. a. Based on your prediction in Question 1, do you believe that your sample came from the hypothesized population or from one whose mean is larger than 10? How did you decide?

   Sample Answers: Students whose sample mean was near 10 are likely to agree that their sample came from the hypothesized population because the sample and the sample mean are "near the center" of the hypothesized population distribution. Students whose sample mean is far to the right (noting that H_a: $\mu > 10$) might suspect that their sample did not come from the hypothesized population.

   > **Teacher Tip:** At some point in this set of questions, you might have to remind students that because H_a: $\mu > 10$, this is a one-sided test, and the only time for concern is when the sample is far to the right.

   b. Show your sample to someone else to see if he or she reaches the same conclusion.

   Sample Answers: Students will most likely agree when samples are centered around 10; they might disagree when samples are clustered or centered far to the right.

   c. Under what conditions would you have reached the conclusion that your sample was not taken from the assumed population H_0: $\mu = 10$?

   Sample Answers: Students should recognize that if a sample mean is far to the right of 10 (i.e., H_a: $\mu > 10$), the sample might not be from the hypothesized population.

   > **Teacher Tip:** Following Question 3c, discuss with the class the criteria they used. Have students view all the class samples. (If you do not have a TI-Navigator, you might, for example, have students walk around all of the desks to view TI-Nspire screens, noting the different samples and how they relate to the hypothesized population.) The discussion should proceed as in the Navigator Tip #1.

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

Move to page 2.2.

The graph in the top work area is a rescaled copy of your Page 2.1. The graph in the bottom work area represents the sampling distribution of $\bar{x}$ for all possible samples of size 5 from the given population.

> **Tech Tip:** After selecting a sample, students might have to click in an empty white space in order to activate the shading in the lower screen.

4. a. Explain why the graph of the sampling distribution of $\bar{x}$ is not exactly the same as the graph of the population.

Sample Answers: While they both have the same mean, 10, the standard deviation of the sampling distribution of $\bar{x}$ is smaller than the standard deviation of the population. If the standard deviation of the population is σ, then the standard deviation of the sample will be $\frac{\sigma}{\sqrt{n}}$ or, in this case, $\frac{\sigma}{\sqrt{5}}$. This makes the spread of the sampling distribution of $\bar{x}$ smaller.

> **Teacher Tip:** Note that the vertical scales are very different in the two graphs. The total area for each is 1, so the lower graph is a lot taller than it appears on the screen.

b. What do you think the vertical line labeled *xc* = 11.1 represents in the graph of the sampling distribution of $\bar{x}$?

Sample Answers: The vertical line bounds the shaded area to the right and passes through the point $\bar{x}$ = 11.1. This would be the mean of a sample that was greater than 90% of the possible sample means from this population.

> **Teacher Tip:** You might want to have students actually do the calculation to show that for μ = 10, σ = 2 and *n* = 5, the 11.1 marks the 90th percentile. This value of the sample mean is sometimes referred to as the "critical" value. This is why the .tns file uses the "*xc*" label.

Samples that are judged as unlikely to have come from the hypothesized population will have sample means that are extreme in the direction of the alternative hypothesis. The term *alpha* (or significance level) is used to quantify what statisticians mean by "extreme enough" and establishes an agreed upon criterion for the meaning of extreme.

5. Look at the mean of your sample in the top graph. Would your sample be judged as unlikely to have come from the hypothesized population for the given alpha level and alternative hypothesis? Why or why not?

Sample Answers: Students with sample means that fall in the shaded area denoted by alpha should indicate that their sample was unlikely to have come from the hypothesized population because this would occur by chance in only 10% of the random samples drawn from that population, and for this situation the researcher has decided that value is "extreme enough."

Teacher Tip: Students should count the points that seem to fall on the vertical line marking the shaded region as being in the region rather than spending time on whether the dot is really in the region.
Following Question 5, use a show of hands to collect the number of students who responded yes to Question 5 and do a quick "ballpark" calculation of the corresponding proportion. Don't point out the direct connection between that proportion and alpha yet—leave that for students to notice as the activity progresses.

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

6. Even though in practice you can usually obtain only one sample, to determine how often samples are classified as unlikely to have come from the hypothesized population, it is necessary to understand the distribution of all samples from that population. In this activity, you still can't look at all samples, but you can look at a lot more than one.
   a. Draw 50 samples, and record the number of times a sample mean falls in the shaded region. What fraction of your 50 samples would be judged unlikely to have come from the hypothesized population?

      Sample Answers: Students should have about five samples whose mean falls in the shaded region.

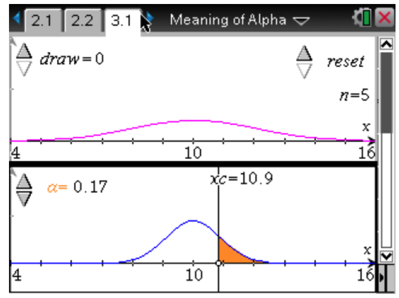
   b. Based on your answers above, what do you think alpha represents? Explain your thinking.

      Sample Answers: Students might begin to articulate that an alpha of 0.1 represents the percent of the sample means that would be "extreme enough" to lead to an "unlikely to have come from the hypothesized population" conclusion.

It is important to understand that ***all*** the samples you have generated (and will generate) in this activity really are random samples taken from the given population. But samples vary. One consequence of that variation is that some samples will appear not to "belong" to a given population even when they really do. The reason for looking at lots of samples from a known population is to begin to understand just how that random variation affects decisions.

Tech Tip: After selecting a sample, students might have to click in an empty white space in order to activate the shading in the lower work area of Page 3.1.

Move to page 3.1, and set alpha to 0.15.



7. a. Explain why the line *xc* is now 10.9.

Sample Answers: The vertical line for *xc* is now marking the 85th percentile rather than the 90th percentile. It shows the border for the top 15% of the area under the distribution.

b. Predict the proportion of samples that would be judged unlikely to have come from the hypothesized population for an alpha of 0.15.

Sample Answers: Students might predict 15% of the samples, or they might not yet have made the connection between the sample means and the shaded area.

> **Teacher Tip:** If students have not yet made the connection between the outcome of the sample means and the shaded area represented by alpha, have them finish the questions below and then come back to their prediction to see if they would like to change it.

c. Draw a sample. What does the point in the lower graph represent?

Sample Answers: The mean of the sample is displayed in the lower graph.

d. Draw 10 samples, and count how many times the sample mean falls in the shaded region.

Sample Answers: Answers will vary from around zero to two of their sample means in the shaded region, with a large variability because of the small number of samples.

Move to page 3.2.

8. Use the arrow to draw 100 samples. Page 3.2 shows the distribution of the 100 sample means from the samples you generated.
   a. What fraction of your 100 samples would be judged unlikely to have come from the hypothesized population based on the given value of alpha?

      Sample Answers: Answers should be around 15/100 or 15%; some students might have as few as 10/100 or as many as 20/100 of their sample means falling in the shaded region. Students might need to estimate the number of dots representing the sample means in the shaded areas. Answers should be expressed in either a fraction or a percent, not just a count, to convey the idea that the results are comparative.

   b. Go back to Page 3.1, and change alpha to 0.1. Return to Page 3.2, and explain what has changed.

      Sample Answers: The set of 100 sample means has remained the same, but the vertical line marking the border for the shaded region (alpha) has moved up to 11.1.

   c. Now, what fraction of your 100 samples would be judged unlikely to have come from the hypothesized population, based on this new alpha?

      Sample Answers: Students should have fewer unlikely samples. Their answers might range from about 7 or 8 to 12 or 13.

> **Teacher Tip:** You might want to have students set a variety of alpha levels and observe how the number of unlikely samples for the hypothesized population changes, depending on the alpha. This can be done using a single set of simulated samples. However, if students want to generate different samples, they should first use the reset arrow to clear Pages 3.1 and 3.2.

TI-Nspire Navigator Opportunity: ***Screen Capture***
See Note 3 at the end of the lesson.

9. Based on your work in Questions 6–8, describe what you believe alpha measures.

   Sample Answers: Alpha measures the probability that you will have a sample mean larger than a predetermined value chosen to define when a sample will be deemed unlikely to have come from the hypothesized population.

When a sample mean indicates the sample is unlikely to have come from the hypothesized population, statisticians would say they "**reject the null hypothesis**." Therefore, the region consisting of all sample means that are "extreme" in the direction of the alternative hypothesis is called the "**rejection region**" for that hypothesis test.

When a sample mean indicates the sample is likely to have come from the hypothesized population, statisticians would say they "**fail to reject the null hypothesis**."

> **Teacher Tip:** Student language might not be quite correct at this stage. Encourage them to think about what they wrote and how it makes the connection between the decision and the alpha level.

10. If a sample mean falls into the rejection region defined by a given alpha value, that sample mean is sometimes called *significant at that alpha level*.
    a. Explain why such a sample mean might be considered significant.

    Sample Answers: If a sample mean falls into the rejection region, it would suggest that you have a sample mean that would occur by chance only as often as specified by that given alpha (i.e., if alpha is 10%, you would get a sample mean that would be as extreme as the boundary for the alpha region in only 1 in 10 outcomes, on average. Thus, to get such a value could be called significant.)

    b. For a sample of size 5 and assuming H_0: $\mu = 10$; H_a: $\mu > 10$, identify each of the following sample means as significant or not and give an explanation for your conclusion. (You may use any of the pages in the .tns file to support your answer.)

    i. $\alpha = 0.1$, sample mean = 12

    Sample Answers: A sample mean of 12 is larger than 11.1, the sample mean that marks the border of the alpha region, so it would be considered significant.

    ii. $\alpha = 0.1$, sample mean = 10.4

    Sample Answers: A sample mean of 10.4 is smaller than 11.1 and is not in the alpha region, so it would not be considered significant.

iii. α = 0.05, sample mean = 12

Sample Answers: A sample mean of 12 is larger than 11.5 and so would fall in the alpha region and be considered significant.

iv. α = 0.03, sample mean = 11

Sample Answers: A sample mean of 11 would be smaller than 11.7 and so would not fall into the alpha region and would not be considered significant.

v. α = 0.01, sample mean = 13

Sample Answers: A sample mean of 13 would fall into the alpha region, as it would be greater than 12.1 and would be considered significant.

c. In which of the cases in part b would you consider the sample mean evidence to reject the null hypothesis? Explain why.

Sample Answers: Any time the sample mean falls into the alpha region, that would be considered evidence to reject the null hypothesis because it would happen by chance with only the specified alpha relative frequency, which the researcher would consider rare. The desired samples are exactly the ones called “significant” in part b so this would happen in i, iii, and v.

11. In all cases, before gathering the necessary sample, the researcher or statistician sets the alpha level that will determine whether the sample mean will be significant, that is, whether the sample mean will lead to a reject or fail to reject conclusion regarding the null hypothesis.
Suppose you were testing a hypothesis using the given alpha. Describe the possible consequences in terms of the null hypothesis.

a. α = 0.45

Sample Answers: You would be deciding to reject the null hypothesis based on evidence that you would get an outcome that falls in the rejection region nearly half of the time based on chance. Your evidence for rejecting the null would be very weak, and you would frequently reject the null when, in fact, you should not have done so.

b. $\alpha = 0.001$

Sample Answers: You would be deciding to reject the null hypothesis based on evidence that you would get an outcome that falls in the rejection region only 1 in 1000 times by chance. You would have very strong evidence that your outcome did not happen by chance. The danger is that you might not reject the null hypothesis when, in fact, you should do so.

Teacher Tip: To follow up on the possible errors made using alpha levels, you might want to use the Statistics Nspired Power Activity.

Teacher Tip: Students should understand that the researcher or statistician sets the alpha level at the outset of the research. As a consequence, one researcher might use an alpha level of 1% while another might use an alpha of 5%. Thus, it is very important for the consumer to understand what the alpha level was when a sample mean is declared significantly different.

Wrap Up

Upon completion of this lesson, teachers should ensure that students are able to understand:

- Alpha level as the probability that you reject the null hypothesis when the null hypothesis is true.
- A sample mean might lead to either a reject or fail to reject conclusion regarding the null hypothesis depending on the alpha level.

TI-Nspire Navigator

Note 1

Question 3, Name of Feature: Screen Capture

Screen Capture the class's samples. For some of the samples, students will all agree those samples came from the hypothesized population. For other samples, i.e. ones to the extreme right, students might disagree. The focus of the discussion should be on the criteria students use to make their decisions about the whether the sample came from the hypothesized population. Direct students toward the criteria that involve measures of center and point out the need for some common agreed-upon standard or criteria.

Note 2

Question 5, Name of Feature: Quick Poll

Use a Quick Poll in TI-Nspire Navigator to collect the responses to Question 5 as explained in the Teacher Tip following Question 5.

Note 3

Question 8, Name of Feature: Screen Capture

A Screen Capture will show each student's set of 100 sample means. Looking at similarities among the distributions and differences in the proportions of samples falling into the shaded region should help students begin to see what alpha actually measures.