Math Objectives

- Students will describe power as the relative frequency of “reject the null” conclusions for a given hypothesis test.
- Students will recognize that when the null hypothesis is true, power is exactly the same as alpha.
- Students will recognize that if the null hypothesis is actually false, power measures the ability to detect an existing difference between truth and hypothesis.
- Students will describe how power is influenced (in predictable directions) by sample size, alpha, and difference between actual and hypothesized parameter values.
- Students will reason abstractly and quantitatively (CCSS Mathematical Practices).

Vocabulary

- alpha level
- power
- critical value
- significance level
- type I error
- type II error

About the Lesson

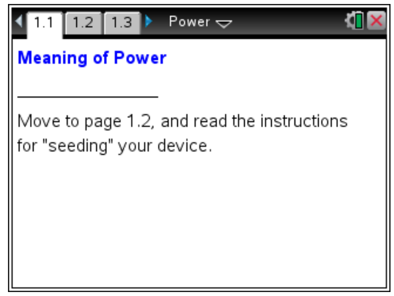
- In this lesson, samples are generated from a population for a particular hypothesis test, leading to the conjecture that the null hypothesis is actually false.
- Students should be familiar with hypothesis testing and understand the meaning of alpha and a type I error.
- As a result, students will:
  - Observe that most of the sample means fall in the rejection region.
  - Examine repeated samples to observe the relative frequency of “reject the null” conclusions—that is, power or alpha—as conditions change.
  - Recognize how the sample size, the true population mean, and the alpha, are related to power.
  - Relate the value of power to beta.

TI-Nspire™ Navigator™

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

Activity Materials

- Compatible TI Technologies: TI-Nspire™ CX Handhelds, TI-Nspire™ Apps for iPad®, TI-Nspire™ Software



Tech Tips:

- This activity includes screen captures taken from the TI-Nspire CX handheld. It is also appropriate for use with the TI-Nspire family of products including TI-Nspire software and TI-Nspire App. Slight variations to these directions may be required if using other technologies besides the handheld.
- Watch for additional Tech Tips throughout the activity for the specific technology you are using.
- Access free tutorials at http://education.ti.com/calculators/pd/US/Online-Learning/Tutorials

Lesson Files:

Student Activity

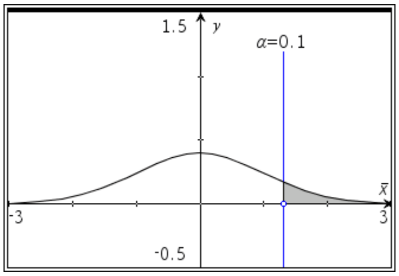
- Power_Student.pdf
- Power_Student.doc

TI-Nspire document

- Power.tns

Discussion Points and Possible Answers

Teacher Tip: Students should not open the .tns file until they have completed Question 1.



Consider the following hypotheses:

H_0: $\mu = 0$

H_a: $\mu > 0$

$\alpha = 0.1$

1. The graph to the right shows a normal curve with $\mu = 0$ and $\sigma = 1$ representing a sampling distribution of sample means taken from a population for which the null hypothesis stated above is true.

   a. What do the values of $\bar{x}$ that fall into the shaded region represent?

   Sample Answers: The shaded region in the graph represents the means for which the null hypothesis would be rejected.

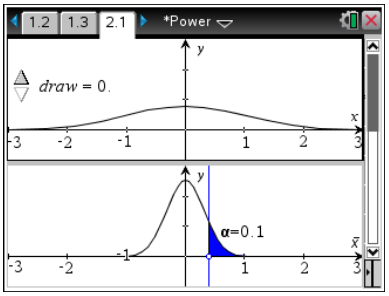
   b. Based on your response in part a and the graph shown, how often would you reject the null hypothesis if the null is in fact true?

   Sample Answers: If the null hypothesis is true, it would be rejected for 10% of the samples drawn from the population.

Tech Tip: Page 1.2 gives instructions on how to seed the random number generator of the TI-Nspire. Page 1.3 is a *Calculator* page for the seeding process. Carrying out this step will prevent everyone from generating identical data. Since some students might have cleared their memory or their handheld is new, completing this process will ensure that students will generate different data. (Syntax: RandSeed #, where # should be a number unique to each student.)

Tech Tip: To enter their random seed value on Page 1.3, students will need to tap the space immediately after the RandSeed command to have the keyboard appear. (Syntax: RandSeed #, where # should be a number unique to each student.) Carrying out this step will prevent everyone from generating identical data. Since some students might have cleared their memory or their handheld is new, completing this process will ensure that students will generate different data.

Move to page 2.1.



This page continues your thinking about the hypothesis test posed in Question 1. Each time you select the top (▲) arrow on Page 2.1, you generate a random sample of size 10. The bottom graph shows the distribution of sample means if the null hypothesis is true. The top graph is the population corresponding to a true null hypothesis. The elements of each sample are displayed as points on the *x*-axis of the top graph. The vertical line represents the mean of the sample. A rejection region corresponding to an alpha of 0.1 is shaded.

Teacher Tip: Students might have to press in a white space on the bottom graph to make the shading for the alpha region show up.

2. a. What does α represent?

Sample Answers: α is the probability that you reject the null hypothesis when the null hypothesis is true.

b. If you drew a sample from this population, where would you expect the sample mean to fall in the bottom graph?

Sample Answers: Most of the sample means will fall to the left of the shaded region in the bottom graph because the assumed mean is 0; and if the null hypothesis is true, only about 10% of the sample means will be in that region.

c. Predict the proportion of samples that will have means that fall in the shaded region.

Sample Answers: 0.1, or 10%, of the samples are likely to have means that fall in the shaded region.

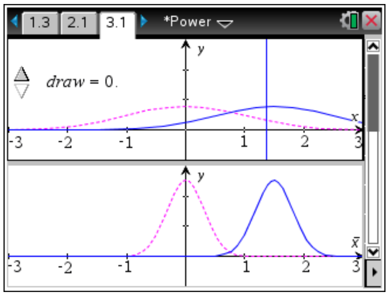
d. Use the arrow (▲) on Page 2.1 to draw 25 samples of size 10. Record what you notice about the pattern of sample means in relation to the shaded region on the bottom graph.

Sample Answers: The mean on the top graph seems to always be to the right of the shaded region shown on the bottom graph.

e. What do your observations in part d tell you about the predictions you made in parts b and c?

Sample Answers: Students should notice that the sample means seem inconsistent with their predictions, and many are falling outside of the sampling distribution—that is, on the bottom screen. They might conjecture that the sample means come from a different population than the distribution on the top.

Move to page 3.1.



The top screen on Page 3.1 represents two populations. The dotted curve centered at 0 represents the null hypothesis. The up arrow (▲) will draw a sample of size 10 from the population you are studying. The bottom screen will accumulate the sample means from those draws. Use the top arrow to generate at least 25 different samples.

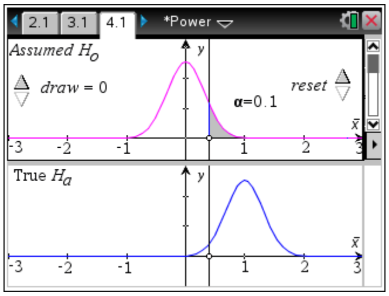
3. a. Which population seems to be the one from which you are sampling?

Sample Answers: Most are accumulating under the solid curve, and only a few or none are under the dotted curve, so the samples would seem to come from the population represented by the solid curve.

b. What does your response to part a suggest about the null hypothesis? Explain your reasoning.

Sample Answers: The null hypothesis is probably false. At this point, students should realize that if the null hypothesis were true, more of the sample means would be accumulating under the dotted curve instead of under the solid curve.

Move to page 4.1.



If the top graph does not have a shaded region, select the top screen so a shaded region appears.

4. In the top graph, α is 0.1.

a. What do $\bar{x}$ values in the shaded region indicate about the hypothesis test being carried out?

Sample Answers: The shaded region represents the set of sample means that lead to rejecting the null hypothesis. If the null hypothesis is TRUE, then the sample means for 10% of all samples will fall into this region by chance, which will lead to rejecting the null hypothesis (an incorrect decision).

b. α is the area of the "rejection region" for the null hypothesis when it is true. Predict the proportion of the sample means you think will fall in the rejection region.

Sample Answers: About 10% of the sample means should fall in the rejection region.

c. Draw one sample by selecting the up arrow (▲) on the top screen. Based on the displayed mean of your sample, what decision (reject or fail to reject) would you make regarding the null hypothesis?

Sample Answers: In most cases, the students' sample mean will fall in the rejection region, and the decision would be to reject the null hypothesis.

d. Draw 25 samples, one at a time. For each sample, tally the decision you make to reject or fail to reject the null hypothesis.

Sample Answers: For most of the samples, you reject the null hypothesis (for example, 24 or 25 of the 25 samples).

e. Are your results from parts c and d consistent with your prediction in part b? Explain what might be happening.

Sample Answers: The results are not consistent with my predictions. Far more samples (nearly all of them) fall in the rejection region. Since alpha measures the probability of rejecting the null hypothesis when the null is true, a rejection rate higher than alpha signals the sampling procedure is flawed or the null hypothesis is false.

5. The bottom curve on Page 4.1 represents the sampling distribution of sample means from a population whose mean is 1. The samples that were drawn in Question 4 actually do come from this population.

a. Given that fact, what decision should you make about the null hypothesis? Explain your reasoning.

Sample Answers: Since the null hypothesis assumed $\mu = 0$, that null is false. Therefore, rejecting the null hypothesis is a correct decision.

b. Use the up arrow (▲) to generate 25 more samples, for a total of 50 samples. What proportion of those samples lead to a rejection of the null hypothesis?

Sample Answers: In 50 samples, I found six sample means that led to a failure to reject; therefore, 44 sample means led to rejection, so my proportion was 0.88.

c. The probability of rejecting the null hypothesis is known as **power**. Thus, *power* is the proportion of the set of all samples' means that fall in the rejection region. Your answer in part b approximates power. To improve that approximation for this situation, draw 50 more samples, making a total of 100. What proportion of all 100 of your sample means lead to the conclusion to reject the null hypothesis?

Sample Answers: Most of the samples will lead to a decision to reject the null hypothesis, so students should respond with "all of them or 100%" or “96 out of the 100 or 96%.”

> **Teacher Tip:** Students should give proportions (percents, decimals, or a fraction) in stating their answer in order to get used to the standard notation for power.

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

6. You studied type I errors earlier. A type II error is failing to reject the null hypothesis when it is false. The probability of committing a type II error is noted as beta, β.

a. What is a type I error?

Sample Answers: A type I error is rejecting the null hypothesis when it is true.

b. For the 100 samples you generated in Question 5, fill out the chart.

Sample Answers:

Decision	Probability	Type of error
Reject null if it had been true	**0.1**	**Type I**
Reject null when false	**0.96**	**not an error**
Failing to reject null if it had been true	**0.9**	**not an error**
Failing to reject null when false	**0.04**	**Type II**

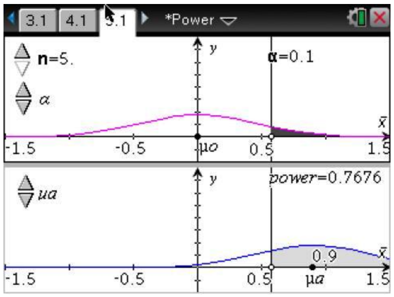
> **Teacher Tip:** Emphasize with students that in this activity, they happen to know that the null is false. Therefore, the two probabilities labeled “if it had been true” are really hypothetical probabilities. They cannot be obtained by simulation within this activity since their underlying hypothesis is false. Similarly, since the null really IS false, the probabilities labeled “when the null is false” CAN be obtained by simulation here.

c. What factors do you think affect power?

Sample Answers: Students might think that power is a function of alpha but might also realize that the probability of rejection when the null hypothesis was false was much larger than the area specified by alpha in the top graph. They might also recall that the sampling distributions of sample means have different spreads for different sample sizes, so sample size may also affect power.

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

Move to page 5.1.



If the top graph does not have a shaded region, select the top screen so a shaded region appears.

7. On page 5.1, α is 0.1, power is 0.7676, n = 5, and the true population mean is μ = 1.0.

   a. Explain what each of these numbers represents.

   Sample Answers: Power = 0.7676 is the probability of rejecting the null hypothesis when the true population mean is actually 1.0 (not 0, as assumed in the null). Alpha = 0.1 is the probability of rejecting the null hypothesis if it had been true. μ = 1.0 is the true population mean. n = 5 is the size of each sample.

   b. Explain the meaning of the $\bar{x}$ values in the shaded region in each of the top and bottom graphs.

   Sample Answers: Each region represents values of sample means that would lead to a "reject the null" conclusion. The shadings in both graphs represent the same $\bar{x}$ values.

   c. What do the areas in the shaded region in each of the top and bottom graphs represent and how are they connected to power?

   Sample Answers: The regions (as represented by the shaded areas on both $\bar{x}$ axes) are identical, but their **AREAS** are not, because the two sets are obtained by sampling from two different populations. The areas represent the probabilities of rejecting the null hypothesis under two different sets of conditions. The area in the upper graph represents the probability of rejecting the null when the null is actually true. The area in the lower graph represents the probability of rejecting the null when the population mean is really 1.0, something other than what was assumed in the null. Power is the area in the lower graph, not the upper graph, as it is the TRUE probability of rejecting the null.

Teacher Tip: It is important for students to understand that area is a number, and a region is a set of points. They should remember that the area under a distribution from $a < x < b$ represents the probability of getting an outcome in the region bounded by *a* and *b*.

Teacher Tip: It is important in this activity to remember that probabilities of events (reject or fail to reject) are determined by the **TRUE** population, not the null. That means that answers to **ALL** of the probability questions are based on the "lower" graph.

Teacher Tip: Some students might suggest that power could be different if the true population mean were something other than 1.0. They would be correct, something they will verify in Question 8, part c.

8. The up arrow (▲) labeled **n** on this page will change the sample size.

   a. Increase the sample size and observe the reported value for power. How does power change? Explain why this seems reasonable.

   Sample Answers: As sample size increases, power increases. If you have a larger sample size, the sampling distribution of the sample means becomes less spread (the standard deviation is $\frac{s}{\sqrt{n}}$). Thus, samples will tend to be more consistently near the true population mean and away from the hypothesized mean, so you will get more samples indicating that the null hypothesis should be rejected.

   b. Select the up arrow (▲) labeled α. How does power change? Explain why this seems reasonable.

   Sample Answers: As alpha increases, so does power because as alpha gets larger, so does the rejection region. Therefore, there is a better chance of rejecting the null hypothesis, and power is defined as the probability of rejecting the null.

Teacher Tip: This connection between alpha and power is reasonable because making alpha larger leads to a rejection region that completely contains the rejection region associated with the smaller alpha, thereby giving a larger probability of rejecting the null hypothesis. That is, the rejection region for a small alpha is a subset of the rejection region for a larger alpha.

c. In the lower half of the screen, use the slider labeled μ_a to change the value of the true population mean. How does power change? Explain why this seems reasonable.

Sample Answers: The further μ_a is from the null population mean, the higher power is. This is reasonable again because you are more likely to reject the null (a correct decision) since sample means are more likely to fall near the true population mean. In particular, the probability of any particular event (such as finding $\overline{x}$ in the shaded portion of the number line) is determined by the ACTUAL sampling distribution (for the TRUE population), not from the hypothesized population and associated sampling distribution.

d. Revisit your conjecture in Question 6, part c about the factors you think affect power. Has your answer changed? If so, how?

Sample Answers: Students should recognize that changes in sample size, alpha, and true population mean affect power, and they should be able to explain why using their examples and reasoning from calculator problems 4–6.

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

9. a. Describe the possible range of values for power and how you know. Explain why the size of power is important.

Sample Answers: Power is between 0 and 1 because it is a probability. The closer power is to 0, the smaller chance you have of rejecting the null hypothesis and the more likely you would be to make the wrong decision with respect to the null hypothesis if the null were false. The closer power is to 1, the more chance you have to reject a false null hypothesis (a correct decision).

b. The probability of committing a type II error (failing to reject the null hypothesis when it is false) is denoted as β. How is β related to power?

Sample Answers: Students should see that power is $1 - \beta$ because of the definition.

c. How do changes in sample size, alpha, and true population mean affect β? Explain why your answers are reasonable.

Sample Answers: As sample size increases, β will decrease since power increases. Similarly, as alpha increases, β will decrease. This is because increasing alpha increases the probability of rejecting the null, so false nulls will be missed less often, decreasing β. When the population mean is further away from the hypothesized mean, β will also decrease since, again, you are more likely to reject the null hypothesis (more samples will fall in the rejection region) and avoid a type II error.

Teacher Tip: Students should realize that it is ideal to have power be as close to 1 as possible, making β as small as possible. Remember that both α and β represent error rates for conclusions from the hypothesis test. Ideally, we would like both to be 0, allowing the hypothesis-test process to be error free. That, of course, is impossible. If only α and β can be controlled, then there is a clear conflict between which will be lower, since decreasing one increases the other. However, sample size is also likely to be under the control of the experimenter, so it provides a tool for keeping both error rates low. How low is "low enough" depends on the context of the test and is determined by the experimenter based on the consequences of each error type.

Teacher Tip: In practice, the true population mean is not known. (If it were known, there would be no need for a hypothesis test!) So, calculating "the" power of a test is impossible. After all, the value of power depends on the actual difference between the hypothesized mean and the true mean. Thus, it is more correct to talk about the power of a test "against μ_a," that is, against a specific alternative, which might or might not be the true mean. In practice, μ_a would be selected as a threshold value for what the experimenter would deem to be a contextually important difference.

10. Write an explanation of the power of a test and why power is important to a test.

Sample Answers: The power of a test is the probability of rejecting the null hypothesis. The higher the value of power, the better the chance you will make the correct decision regarding the null hypothesis if it is false.

Wrap Up

Upon completion of this activity, students should be able to understand:

- Power as the relative frequency of “reject the null” conclusions for a given hypothesis test.
- When the null hypothesis is true, power is exactly the same as alpha.
- How power is influenced (in predictable directions) by sample size, alpha, and the difference between actual and hypothesized parameter values.

TI-Nspire Navigator

Note 1

Question 5, Name of Feature: Quick Poll

Use Quick Poll to check student responses to Question 5 part b. Most of the responses should be similar; discuss with students what this might indicate about the null hypothesis.

Note 2

Question 6, Name of Feature: Quick Poll

Use Quick Poll for Question 6 part c, where students indicate the factors that might affect power. The discussion can focus on why the elements they identified would affect power.

Note 3

Question 8, Name of Feature: Class Capture

A Class Capture of Question 8 part a, b, or c would show the effect of the different settings on the rejection region for the null hypothesis. Students could explain how to interpret the different outcomes and to identify the outcomes that seem to be very similar.