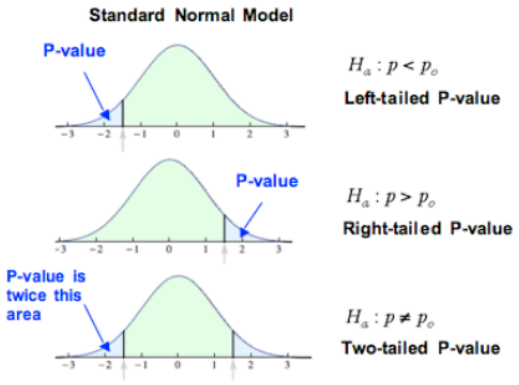
Learning Goals:

- Recognize when a situation calls for testing a hypothesis about a population proportion.
- Conduct a hypothesis test for a population proportion. State a conclusion in context.

Introduction:
In this activity we continue our work with hypothesis testing. This first example is what we call a "two-sided" test. We use a two-sided test when our claim is that there is a change, as opposed to an increase or a decrease, in a population proportion. In this situation, the alternative hypothesis has a "≠" symbol, instead of a "greater than" or "less than" symbol.



Example: CCSSE is the Community College Survey of Student Engagement. CCSSE was launched in 2001, with the intention of producing new information about community college quality and performance. In the CCSSE student survey data is used to help community colleges set benchmarks to improve student learning and retention.

According to the 2016 CCSSE data from about 430,000 students nationwide, about half of the students reported that they "often" or "very often" prepared two or more drafts of a paper or assignment before turning it in.

Are the results similar at LMC? Specifically, let's test the claim that LMC is different. Use a 5% level of significance.

Step 1: Determine the hypotheses.

Ho: p = 0.50
Ha: p ≠ 0.50

Step 2: Collect the data and report the sample results.

Suppose that in a random sample of 200 LMC students, 56% report that they "often" or "very often" prepared two or more drafts of a paper or assignment before turning it in.

Step 3: Assess the data.

Are these results statistically significant? Or is the 6% difference between 56% and the hypothesized 50% just due to the variability we expect to see when we collect random samples? These questions can be answered with a P-value.

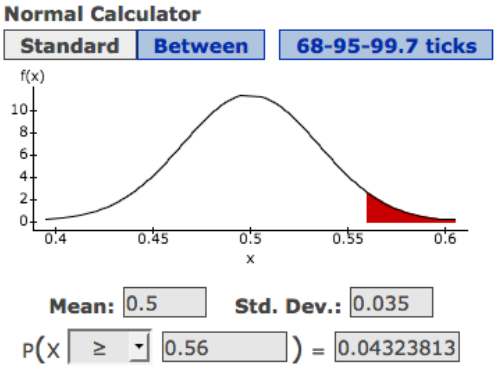
- Verify that the conditions are met for the use of a normal model for the distribution of sample proportions.

- What are the mean and standard deviation of the distribution of sample proportions in this situation?

- Are the results statistically significant?

  Use the StatCrunch Normal Calculator to find the P-value by doubling the P-value associated with a one-sided test.

  What is the P-value?



Step 4: State a conclusion

Group work

1) According to the 2016 CCSSE data from about 430,000 community college students nationwide, about 13.5% of students reported that they "often" or "very often" come to class without completing readings or assignments.

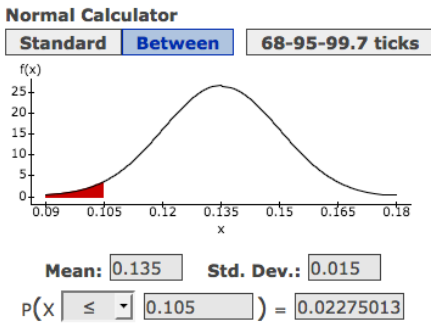
   Are the results similar at community colleges in California? Specifically, let's test the claim that California students are different. Use a 5% level of significance.

   Suppose that the CCSSE is given in California to a random sample of 500 students and 10.5% report that they "often" or "very often" come to class without completing readings or assignments.

   a) State the hypotheses and write a sentence to explain what p represents.

   b) Verify that the normal model is a good fit for the distribution of sample proportions.

   c) Verify that the mean and standard deviation of the distribution of sample proportions is as shown in the Normal Calculator. Explain or show the formula(s) with numbers plugged in.



   d) Use the image of the Normal Calculator to find the P-value.

   e) State your conclusion.

2) Do you think your conclusion in the previous problem would change if the 10.5% came from a random sample of 100 students instead of a random sample of 500 students? Why or why not?

3) A 2014 Fox News Poll reported that 20% of voters is tattooed. Has the percentage changed this year?

Suppose that in a survey of 1000 voters this year, 22.5% are tattooed.

We use StatCrunch to run the hypothesis test (Stat, Proportion Stats, One sample, With summary)

One sample proportion hypothesis test:
p : Proportion of successes
H_0 : p = 0.2
H_A : p ≠ 0.2

Hypothesis test results:

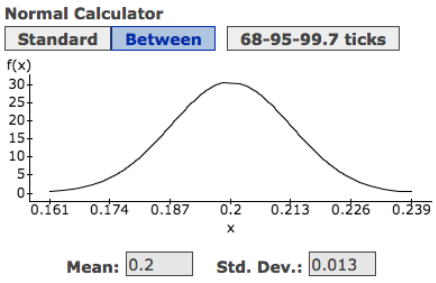
Proportion	Count	Total	Sample Prop.	Std. Err.	Z-Stat	P-value
p	225	1000	0.225	0.012649111	1.9764235	0.0481

a) Write a sentence to explain what p represents in the hypotheses.

b) What wording in the problem statement suggests that the alternative hypothesis should be "≠"?

c) The yardstick in the normal curve is standard error. Shade the total area under the normal curve that represents the P-value.

d) Does the StatCrunch printout give the P-value or do we have to double this number? How do you know?



e) What do we need to know about this sample in order to conduct a hypothesis test?

4) According to the 2016 CCSSE data from about 430,000 community college students nationwide, about 26% of students reported that they never made a class presentation.

The Communications Department at a local community college is trying to encourage faculty in other departments to require class presentations. They think that class presentations are less prevalent at their college than at other colleges.

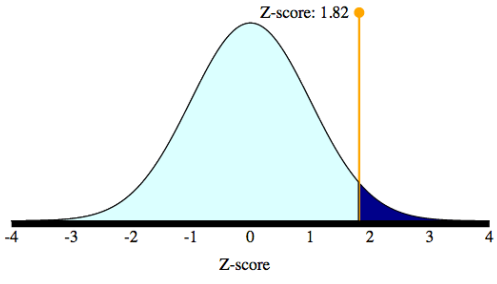
To study this issue they survey a random sample of 100 students and find that 34% have never made a class presentation. This sample supports what the Communications Department thinks is true, but is this 8 percentage point difference significant? We have to conduct a hypothesis test to know.

a) State the hypothesis. Write a sentence to describe what p represents.

b) Verify that the normal model is a good way to estimate the P-value.

c) Verify that the standard error is approximately 0.044. Show the formula with correct numbers plugged in. What does this number represent?

d) This time we used the OLI Normal Distribution Calculator, which is based on z-scores. Verify that the z-score for the sample is about 1.82. Show the formula with correct numbers plugged in.



The area to the left of the Z value is: 0.9656

The area to the right of the Z value is: 0.0344

e) What is the P-value?

f) What conclusion can we draw?