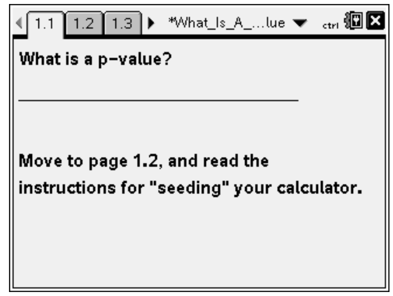
What is a p-value?
Student Activity

Name ________________
Class ________________

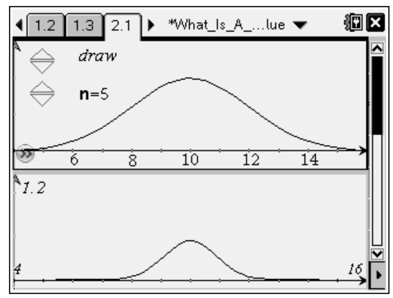
Open the TI-Nspire document *What_Is_A_P-value*.



Suppose you draw a sample from a population you assume has a given mean and calculate a sample mean from your sample. How likely are you to draw a sample that will have a mean at least as extreme as your observed sample mean? This activity will help you think about this question.

Move to page 2.1.

Press ctrl ▶ and ctrl ◀ to navigate through the lesson.



Consider the following hypothesis test scenario:

H_0: $\mu = 10$ and H_a: $\mu > 10$

The top screen of page 2.1 of the tns file represents a population whose mean is 10, as assumed in the null hypothesis.

1. Use the up arrow (▲) to draw a random sample of size 5 ($n = 5$) from this population.
    a. What do the dots in the top graph represent?

    b. What does the white dot in the bottom graph represent?

    c. Estimate the values of the elements in the sample and give the sample mean.

2. a. Considering the alternative hypothesis stated above, does it seem likely that your sample came from the hypothesized population or from a population whose mean is considerably larger than 10? Explain your reasoning.

    b. Estimate the probability that future sample means will fall to the right of this sample mean.

Use the random sample you have for Question 1 to answer the next set of questions.

The term *p-value* is the probability that you would get a sample mean at least as extreme as the one from the random sample you drew if the null hypothesis is correct. The shaded area in the lower screen indicates the p-value.

3. a. Explain why this area represents a probability.

   b. Interpret the *p*-value you have on your screen.

   c. Based on your observed *p*-value, does it seem likely that your sample came from a population with $\bar{x}$ = 10? Why or why not?

Move to 2.2, without changing the *p*-value you found above.

4. a. Without actually drawing additional samples, describe the sampling distribution of the sample means from the samples you would expect to get.

   b. Now use the up arrow (▲) to draw more samples of size 5 from the given population. Explain what the values plotted on the horizontal axis represent.

   c. How does the simulated sampling distribution of the sample means compare to your predictions in part a? Explain any differences.

   d. From your simulation of sample means, estimate the likelihood of getting a sample from the given population so that the new sample mean is at least as extreme as the original observed sample mean.

5. Use ctrl tab to move to the line below the screen with the graph. Type **reset()** in the box and press enter. (These instructions are also on **page 2.3** of the tns file.) Resetting will allow you to simulate another sampling distribution of sample means from samples of size 5.

   a. Before generating new samples, predict how the new distribution of sample means will compare to the one you generated for Question 4.

   b. Remember that the vertical line labeled $\bar{x}$ represents your original sample mean. Predict the proportion of future samples that will have means to the right of that sample mean.

   c. Generate 100 sample means. How well does their distribution match your predictions in parts a and b? Explain any differences.

Return to page 2.1.

6. Change the sample size to 10 (n = 10). Draw another random sample.
   a. What are the sample mean and the corresponding p-value for this random sample?

   b. Describe what your p-value means.

   c. Compare your p-value with that of other students in your class. Why do these values differ?

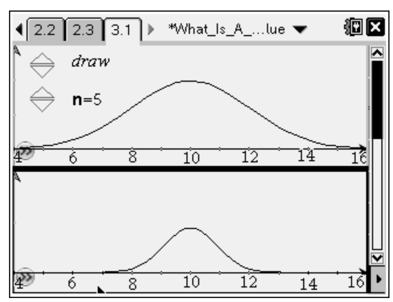
   d. Sketch the simulated sampling distribution of sample means you might expect to get if you were able to draw 100 samples rather than just one.

   e. Predict the proportion of future samples that will have means to the right of your displayed sample mean.

Move to page 2.2, using your new *p*-value.

7. Draw 100 samples of size n = 10 and compare this simulated sampling distribution of sample means to your prediction in Question 6 parts c and d.

Move to page 3.1.



8. Assume H_0: $\mu = 10$ and H_a: $\mu > 10$.
   a. Draw samples from the population for n = 5 until you find a p-value less than 0.1. What are your sample mean and the corresponding p-value?

   b. Change the sample size to n = 10, and draw samples until you get a sample mean close to the sample mean you found in part a. What is the corresponding p-value? Repeat the process for sample size n = 15.

   c. Jordan found an observed sample mean for a sample size of n = 5 and claimed that a sample size of n = 20 with the same sample mean would have the same p-value. Explain whether you agree or disagree with Jordan and why.

9. All of the work in this activity has assumed that H_0: $\mu = 10$ and H_a: $\mu > 10$. How would your thinking change if H_a: $\mu < 10$?

10. In earlier work, you studied alpha levels. Describe similarities and differences between alpha levels and *p*-values.

11. Identify each statement as true or false, and give a reason for your choice.
    a. The decision to reject or not reject the null hypothesis is based on the size of the *p*-value.

    b. If the null hypothesis is true, a *p*-value of 0.05 would mean that on average one out of 20 samples would result in a mean at least as big as that observed in our random sample just by chance.

    c. If a *p*-value is 0.04 and H_a is a "greater than" hypothesis, there is a 96% chance that the sample mean you observed is from a sample from a population with a mean equal to the null hypothesis.

    d. Small *p*-values suggest that the null hypothesis is unlikely to be true.

    e. In working with *p*-values, you always begin with a set significance level.

    f. An observed sample mean can be significant at one alpha level but not significant at another.