**Statistics 200: First Lab Activity for Section 4.3**

**At this time in the semester the students finally have all the tools they need to perform a complete hypothesis test. As you’ll see in lecture, we teach them that there are five steps/parts:**

1. **state the null and alternative hypotheses.**
2. **determine value of observed statistic**
3. **Find the p-value (using a randomization distribution)**
4. **make a generic conclusion about H0 (reject, fail to reject)**
5. **Make a conclusion in context describing the strength of the evidence for Ha and referring back to the question of interest.**

**The main goal of this lab is to get students testing hypotheses. We start by providing them with lots of guidance about what to do, then gradually remove the structure and expect them to know how to complete the tests on their own.**

**Note: Students typically have no problem with the generic conclusion (reject or fail to reject by comparing p-value to 0.05). They DO have problems with the conclusion in context. They need to understand if the sample provides some evidence, strong evidence, very strong evidence, etc. There is a good graphic on page 295 of the text that shows informal strengths of evidence against H0 and for Ha based on a p-value. Refer them to lecture notes if possible, and this figure if not possible.**

**This is another lecture that I can recommend attending.**

**Determining Statistical Significance - Learning objectives:**

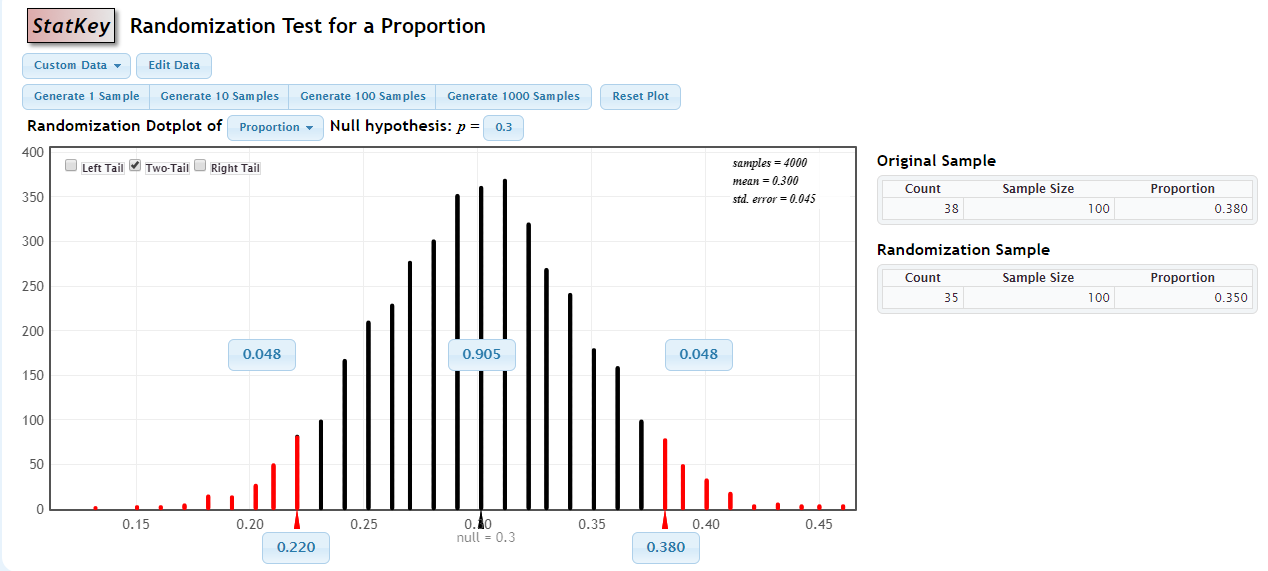
* Recognize that smaller p-values give stronger evidence in support of the alternative hypothesis
* Demonstrate an understanding of the concept of statistical significance
* Make a formal decision in a hypothesis test by comparing a p-value to the significance level
* State the conclusion to a hypothesis test in context
* Make a less formal decision that reflects the strength of evidence in a p-value
* Conduct a hypothesis test for a variety of situations

**Activity 1**:  ***p-values, evidence for the alternative, and formal decisions.***

Below are three different hypothesis tests about population proportions. For each test, use StatKey and the information given to calculate the appropriate p-value and make the correct conclusion.

1. **H0: p = 0.3 vs Ha:  p ≠ 0.3.** In their survey they had a count of 38 using a sample size n=100.
2. What is p-hat for this sample?
3. Using StatKey, generate a randomization distribution using at least 4000 samples. Remember to select Edit Data to input sample information, and to edit the null hypothesis.

What is the p-value using this randomization distribution? **p-value = 0.096**

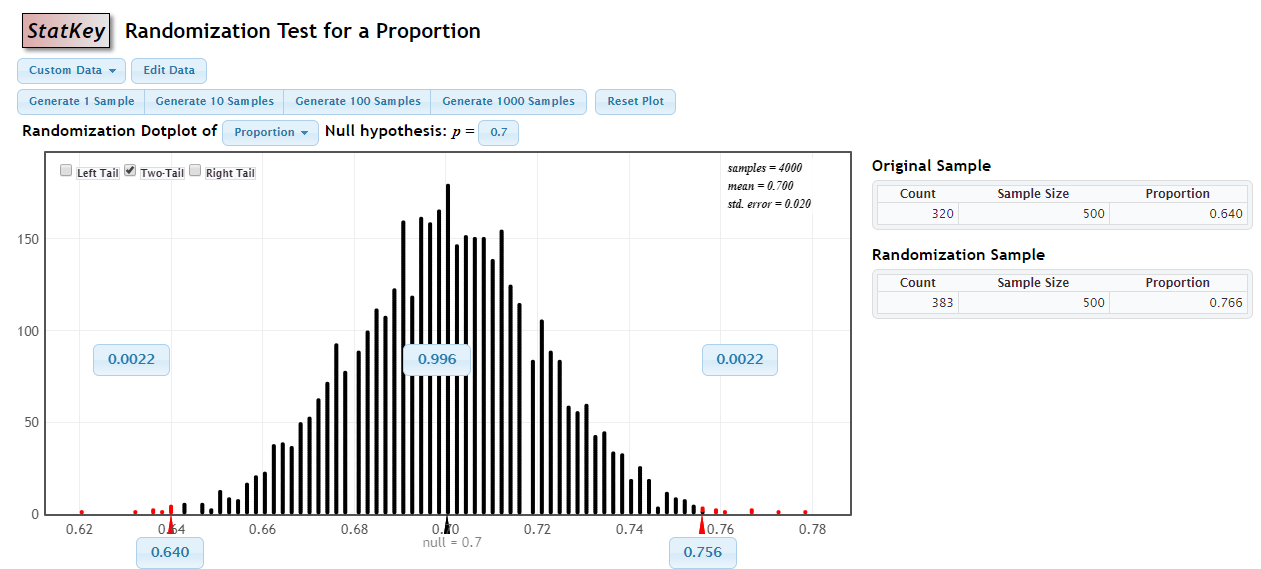


1. At a significance level of 0.05, what is the conclusion for this hypothesis test?

**We cannot reject the null hypothesis at a significance level of 0.05.**

1. **H0: p = 0.7 vs Ha:  p ≠ 0.7.** In their survey they had a count of 320 using a sample size n=500.
2. What is p-hat for this sample?
3. Using StatKey, generate a randomization distribution using at least 4000 samples. Remember to select Edit Data to input sample information, and to edit the null hypothesis.

What is the p-value using this randomization distribution? **p-value = 0.0044**

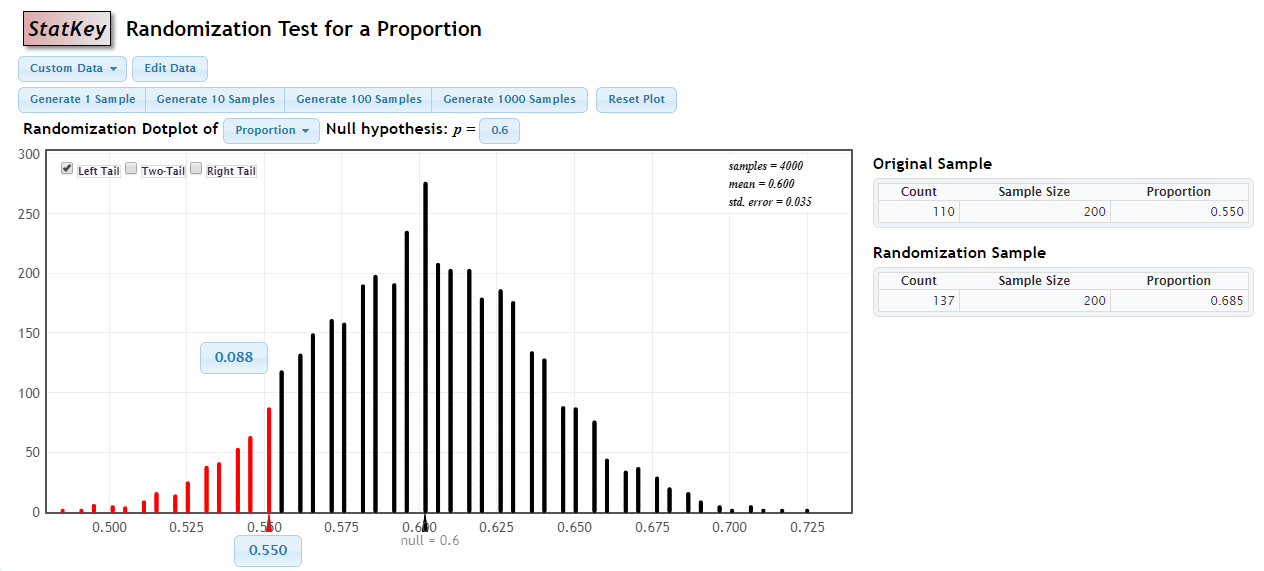


1. At a significance level of 0.05, what is the conclusion for this hypothesis test?

**We can reject the null hypothesis at a significance level of 0.05. Thus, we accept the alternative hypothesis at a significance level of 0.05.**

1. **H0: p = 0.6 vs Ha:  p < 0.6.**  In their survey they had a count of 110 using a sample size n=200.
2. What is p-hat for this sample?
3. Using StatKey, generate a randomization distribution using at least 4000 samples. Remember to select Edit Data to input sample information, and to edit the null hypothesis.

What is the p-value using this randomization distribution? **p-value = 0.088**



1. At a significance level of 0.05, what is the conclusion for this hypothesis test? **We cannot reject the null hypothesis at a significance level of 0.05.**
2. Considering your p-values from questions 1 – 3 above, which of the statements below is **true**?
3. The p-value for Question 2 indicates the most evidence in support of the corresponding alternative hypotheses.

**TRUE. The least p-value gives the most evidence for the alternative, question 2 has the lowest p-value.**

1. For Question 1 we accept the null hypothesis at a significance level of 0.05 and of 0.10.

**FALSE. The p-value is greater than 0.05, but less than 0.10. So we would reject the null at a significance level of 0.10, but not at a level of 0.05.**

1. For Question 2, we would reject the null at a significance level of 0.10, but not at a significance level of 0.05.

**FALSE. Since the p-value is less than 0.05, we would reject the null at both levels of significance.**

1. For Question 3, the conclusion would change if it were a two-tailed test with the same sample statistic.

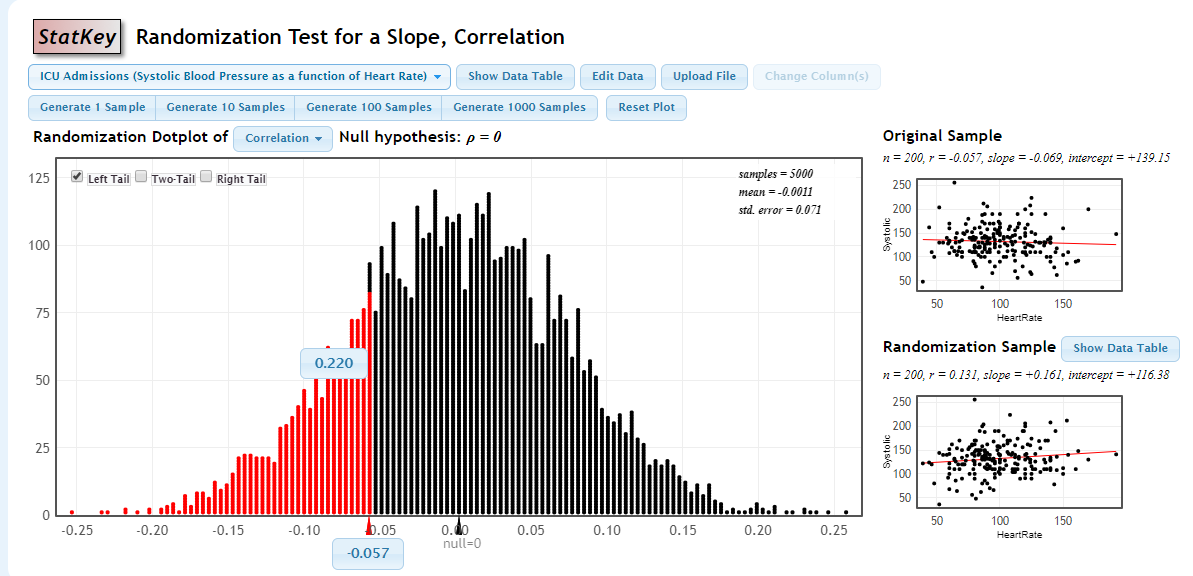
**FALSE. For question three, the p-value would be doubled if it were a two-tailed test. Either way, we would still fail to reject the null, since the one-tailed p-value is more than 0.05.**

**Activity 2: *Hypothesis test from start to finish***

Is there evidence of a negative correlation between systolic blood pressure and heart rate? In a sample of 200 patients, we found a sample correlation of -0.057.

1. State the hypotheses of interest.
2. What is the notation and value of the sample statistic**?**
3. Use StatKey to generate a randomization distribution for these hypotheses. Use the data set ‘ICU Admissions’ available on StatKey.

What is the p-value? **The p-value is 0.220.**



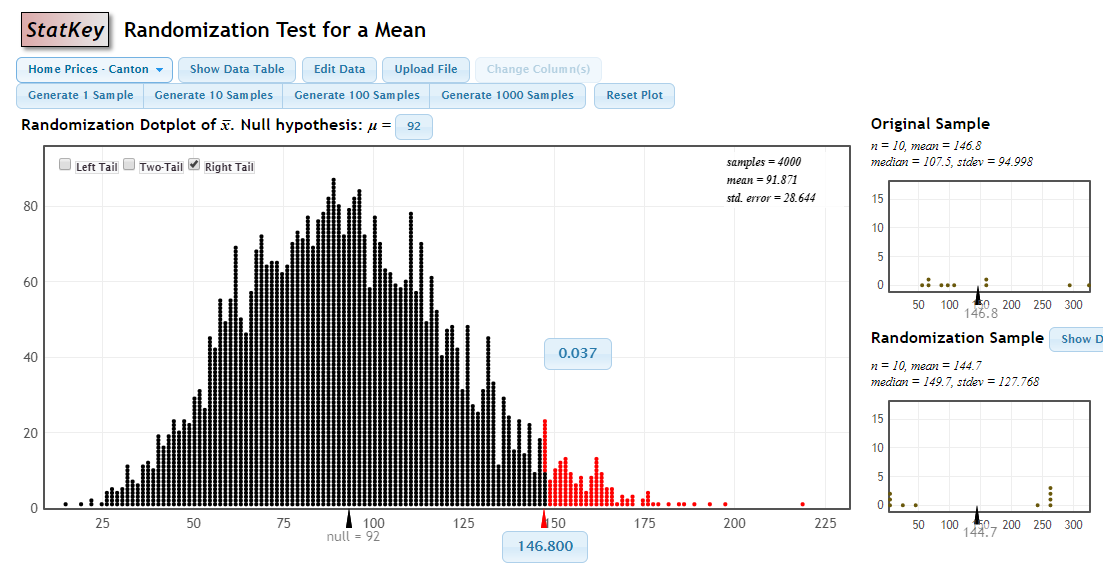
1. Which two of the options below are **correct** interpretations of the p-value (insert your calculated p-value from (3) in the blank):
   1. **If there is no linear association between systolic blood pressure and heart rate in the population, the chance of seeing a sample correlation of -0.057 or less is 0.220.**
   2. If there is a negative linear association between systolic blood pressure and heart rate, the chance of seeing a sample correlation of -0.057 or less is 0.220.
   3. **If there is no linear association between systolic blood pressure and heart rate in the population, the proportion of samples with a sample correlation of -0.057 or less is 0.220.**
   4. If there is a negative linear association between systolic blood pressure and heart rate, the proportion of samples with a sample correlation of -0.057 or less is 0.220.
2. What is the formal conclusion at a significance level of 0.05? **Because the p-value of 0.220 is greater than the significance level of 0.05, we cannot reject the null hypothesis that at a significance level of 0.05.**
3. What is the conclusion in context? **The sample does not contain enough evidence to support the claim that there is a negative correlation between blood pressure and heart rate.**
4. Based on the p-value you calculated in part 3, what would the correct formal conclusion be if we were testing a two-sided hypothesis with the same sample?
5. Reject H0 and accept Ha
6. Reject Ha and accept H0
7. **Fail to reject H0**
8. Fail to reject Ha

**Activity 3*: Hypothesis test from start to finish***

According to Zillow, the median home price in Canton, OH is 92 thousand dollars. A real estate developer believes that the *mean* home price is larger. We are to use a sample of n=10 home sales to see if there is evidence that the mean home price is larger than 92 thousand dollars. The sample mean is 146.8 thousand dollars.

1. State the hypotheses of interest.
2. What is the notation and value of the sample statistic?

3. Use StatKey to generate a randomization distribution for these hypotheses. Use the data set ‘Home Prices - Canton’. You may need to change the null value. What is the p-value? **The p-value is 0.037.**



1. Select the option that gives the correct interpretation:

The p-value from part 3 gives the chance of seeing a sample mean of \_\_\_\_\_\_\_\_\_\_\_ or any value \_\_\_\_\_\_\_\_\_\_ , assuming the population mean home price in Canton, Ohio is truly \_\_\_\_\_\_\_\_.

* 1. 146.8, smaller, 92
  2. **146.8, larger, 92**
  3. 92, smaller, 146.8
  4. 92, larger, 146.8

1. What is the formal conclusion using a significance level of 0.05? **Since the p-value of 0.037 is less than 0.05, we can reject the null hypothesis at a significance level of 0.05.**
2. What is the conclusion in context? **There is enough evidence in this sample to conclude that the mean price of homes in Canton, OH is greater than $92k.**
3. What would the p-value and conclusion be if we decided to make the alternative hypothesis two-tailed?  **The p-value would be 0.074 if it was a two-sided alternative hypothesis, and we would fail to reject the null hypothesis in that situation.**
4. Select everything that would result in a p-value *larger* than that in part 3 above:
5. **Making the alternative hypothesis two-sided.**
6. **Changing the null value from 92 to 96**
7. Changing the null value from 92 to 90

(note – you can do these things in StatKey to see what happens, but you should be able to answer this question without it)

1. Did the sample from Activity 2 or Activity 3 provide more evidence for the corresponding alternative hypothesis?
2. Activity 2
3. **Activity 3**
4. Both provided the same amount of evidence