Errors and Power Corey Andreasen, candreasen@sasd.net

The Situation:

Suppose we are developing a drug to address a particular type of skin rash. This condition goes away within a month on its own in about 40% of cases. We have a topical medication we hope will help more people be cured.

One of these statements describes the reality:

1. The new medication does not make any difference in cure rates.

2. The new medication does make a difference in cure rates.

**Q1**. One of the statements above corresponds to the null hypothesis. Which one is it?

**Q2**. Write the hypotheses in symbols.

The Test:

We plan to do a randomized experiment, giving one group of people the medication and the other a placebo. When we do a significance test on our data, we will *conclude* one of the following:

1. We don’t have evidence that new medication makes a difference in cure rates.

2. We do have evidence that new medication makes a difference in cure rates.

Unfortunately, the reality and our conclusion may not match.

We will imagine an experiment with 100 people in each group. Open the file **Errors\_and\_Power.ftm**. In this activity, *p* is the actual underlying proportion of people who get better using the medication and *p*0 is the proportion of people that get better without the medication, so it is our hypothesized value of *p*.

There are four sliders to control different details about our simulation. p0 controls the proportion of people using our medication that will get better under the null hypothesis. In our example, that is set at 0.4.

The second slider, p, controls the *actual* proportion of people using the medication that will get better.

The third slider, *n*, controls the number of subjects in each group. For simplicity, we will make both groups the same size.

The fourth slider, *alpha*, lets you choose the significance level for your significance test.

**Q3**. Examine the formulas for each attribute in the table. Explain what each formula does.

Possible Reality #1: The new medication does not make a difference in cure rates.

**Q4**. If we use the new medication, in what percent of people would we expect to see the condition go away in a month if the medication does not work better than nothing?

Generate 1000 cases. (Choose **Collection|New cases**, then type 1000 in the dialog box.) This represents repeating the experiment 1000 times.

**Q5**. What do the two values shown by the vertical lines on the graph indicate?

Drag a **New Summary Table** from the object shelf. Drag the attribute *Difference* to the table and create two formulas. First, double-click the formula for the mean and change it to

 . Then choose **Summary|Add Formula**, and use the formula . (*Bound* is a name given to the value of the positively valued vertical bar.) If you have moved the sliders, reset them to their original values: p0 = 0.4, p = 0.4, n = 50, and alpha = 0.05.

**Q6**. a. What is the probability that a particular experiment results in a difference between the vertical bars?

 b. What decision would you make if that happens? Does this decision match reality?

**Q7**. a. What is the probability that a particular experiment results in a difference outside the vertical bars?

 b. What decision would you make if that happens? Does this decision match reality?

The situation described in **Q7** is referred to as a *Type I Error*.

**Q8**. Describe in words the definition of a Type I Error. Explain what this has to do with **.

**Q9.** What is the consequence of making a Type I Error in this situation? Specifically, what does that mean for potential patients?

Possible Reality #2: The new medication does make a difference in cure rates.

For the moment, assume that 60% of people who use the medication will be cured. To set this up, drag the *p* slider to 0.6.

**Q10**. Describe what the normal curve represents.

**Q11**. Describe what the histogram represents.

**Q12**. a. If *p* = 0.6, what is the approximate probability of an experiment resulting in a difference that falls between the bars?

 b. What decision would you make if that happens? Does this decision match our new reality?

This situation is referred to as a *Type II Error*. The probability of a Type II Error is represented by they symbol ** (Beta).

**Q13**. Describe in words the definition of a Type II Error.

**Q14.** What is the consequence of making a Type II Error in this situation? Again, what does it mean for potential patients?

**Q15**. a. If *p* = 0.6, what is the approximate probability of an experiment resulting in a difference that falls outside the bars?

 b. What decision would you make if that happens? Does this decision match reality?

This probability is called the *power* of the test.

**Q16**. Describe in words the definition of power.

Adjust the different sliders and observe the effects on the power of the test.

**Q17**. In this situation, what does it mean for *p* to change? How is the power affected if *p* changes? What about the probability of a Type II error?

**Q18**. How does changing the sample size affect the power? How does it affect the probability of a Type II Error?

**Q19**. During what phase of a statistical investigation is power a consideration?

Explore Further:

**Q20**. What effect does changing the significance level have on power? What are the other consequences of making that change?