

Homework 3. Submit by email by 10am on Wednesday, July 25.

As I mentioned in class, you can find interactive websites that do all the work of calculating chi-square probabilities for you once you've plugged in your observed values. Here's one: <http://www.psych.ku.edu/preacher/chisq/chisq.htm>. But it's useful to go through the exercise of doing the test manually at least a few times so that you can develop a sense for what the test is actually doing and when it is appropriate to use it.

In general, **chi-square is a good test to use if:**

- You have enough observations so that all of the values in the (2x2) Expected table are 5 or greater.
- You are working with count data, rather than percentages.
- Your categories are discrete (e.g. either male or female, either 0 or 1).
- Your observations are independent. This isn't the right test to use if you're comparing e.g. 'before' and 'after' values for the *same* set of individuals.

Also keep in mind that the chi-square test itself doesn't 'know' what your hypothesis is. All it will tell you is what the probability is that your observed sample could have been drawn by chance. It's possible that you could get a statistically significant result that goes against your hypothesis – for example, if you thought women would have more /ae/-tensing than men but they actually had much less.

1. For the first part of the assignment, **create your own chi-square calculator** for a 2x2 contingency table using Excel or a comparable program. Use the values from the exercise we did together in class to create and test your calculator.

There are a few different ways you can do this. The functions you'll most likely need are CHIDIST and CHITEST (they're the same in OpenOffice as far as I can tell). Use the Help menu (F1 in Excel) to read about how they work, and then experiment with them.

Your final chi-square calculator should be:

 - (a) user-friendly (it should be clear to the user where to enter their observed data)
 - (b) adaptable to any range of values I might enter into the 2x2 table (in other words, make sure to use references to *cells* where appropriate, rather than entering specific values associated with a particular observation).

2. Next, pick an observation from either our preliminary /ae/-tensing study, the *ing* data from Homework 2, some other study we've read about in class (assuming you can recover all the information you need, i.e. *n*'s), and apply the chi-square test to see if the observed sample is significant or not at the (i) $p < .05$ and (ii) $p < .01$ levels. Go through the following steps:
 - Make sure your observation has the properties in the bulleted list above.
 - Summarize your observation, and state your hypothesis and the null hypothesis.
 - Calculate the chi-square probability with the calculator you made in part 1.
 - Check your result with the online tool cited above.
 - Report your result in a sentence or two.

Depending on what you decide to look at, you may decide to set up a larger contingency table (e.g. 2x3 for Progressive, Noun, and Verbal *-ing* words).

Include this part of the assignment in the same spreadsheet as your calculator. Make sure each part of the assignment is labeled clearly. Email the spreadsheet to me by **10 am on Wednesday, 7/25**.