**Unit 12, Module 32: Chi-Square Test for TWO Way Tables**

In this module, we will learn about how to conduct a Chi-Square Test for INDEPENDENCE. This is the ONLY Chi-Square Test that we will do (there is another about Homogeneity in your textbook, but we don’t have sufficient time to do both).

For a Chi-Square Test for Independence, we start off with the following hypotheses:

 (INDEPENDENCE)



We then run a hypothesis test that is similar to the way we completed a Chi-Square Test for 1 variable. We will demonstrate this by completing the following example.

The following is observed data from a few semesters of statistics surveys:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Feelings about Math | | | | |  |
|  | A | B | C | D | E | **Total** |
| Female | 5 | 21 | 10 | 10 | 10 | 56 |
| Male | 9 | 8 | 9 | 10 | 8 | 44 |
| **Total** | 14 | 29 | 19 | 20 | 18 | 100 |

**Step 1:** First, what should the Hypotheses be for a Test of Independence? Write them here:

Next, we will have to compare our observed results with the **EXPECTED** results. In order to find the **EXPECTED** results, we have to figure out what the results ***should*** be, IF the Null hypothesis were true – meaning, if the two variables “gender” and “feelings about math” were indeed independent. We do this by finding the overall percentage of A’s, B’s, C’s, D’s and E’s. What percent of everyone chose A? (A = LOVE math!) Find all of the necessary percentages:

**Step 2:** Next, we can fill out an EXPECTED chart…that is, a chart that looks very similar to the one presented above, but with all EXPECTED values in it.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Feelings, EXPECTED | | | | |  |
|  | A | B | C | D | E | **Total** |
| Female |  |  |  |  |  | 56 |
| Male |  |  |  |  |  | 44 |
| **Total** | 14 | 29 | 19 | 20 | 18 | 100 |

If 14% of everyone loves math, if gender were independent of ones feelings about math, we would expect that 14% of females would love math and 14% of males would love math. This is how we fill out the table.

**Step 3:** Next, we must calculate Chi-Square. From last time: . To be sure we don’t forget any of the boxes, we can use a blank chart to help us out:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Chi-Square Contribution | | | | |
|  | A | B | C | D | E |
| Female |  |  |  |  |  |
| Male |  |  |  |  |  |

We find  for each observed value, so each box should get filled in. Then we add up all of the boxes to get .

**Step 4:** Now, we need to find the p-value and just like with one-way tables, we need to use technology. This time, finding the degrees of freedom is a little different. When there was just ONE variable, we just found r – 1 (number of categories – 1), but now we have TWO variables. To find the degrees of freedom now, we need to multiply:



Find the degrees of freedom for this problem:

Then we use technology (the same website as last time) and plug in the degrees of freedom and the value of  to find the p-value. You can also draw a picture, if that is helpful to you (do you remember what the distribution of  looks like???)!

**Step 5:** We draw conclusions in the same way as a hypothesis test – do you remember how? What should we conclude?

Checks/Assumptions: What do you need to check to be sure that the results of this test are valid?