**Learning Target** • **Chi-Squared Test for Two-Way Tables**

We can do a significance test for categorical data in a two-way table using a Chi-Squared ($X^{2}) $test. We test to determine if two categorical variables are related to each other. The procedure is the same PHCTAC, but the test statistic is different (instead of z, we will use *X* 2).

Since we will be working with categorical data, let’s review of 2-way tables and segmented bar graphs:

Example 1: Cocaine addicts are often addicted to feeling pleasure. A study was done to determine if giving addicts an antidepressant would help them stay off cocaine. In a 3-year study, 72 addicts who wanted to break their habit were randomly assigned to 3 treatment groups (Desipramine: antidepressant, Lithium: current addiction treatment, and a placebo). Those successful were able to stay off cocaine during the study. The results are recorded in the table below:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Success | Failure | Total |
| Desipramine | 14 | 10 | 24 |
| Lithium | 6 | 18 | 24 |
| Placebo | 4 | 20 | 24 |
| Total | 24 | 48 | 72 |

Create a segmented bar graph of success and failure for each treatment. Does there appear to be a difference in the treatments?

The segmented bar graph gives a good visual, a calculation still needs to be done to determine if differences are significant (lower than α)

The Chi-Square test is done comparing the observed values (actual data) from the expected values. The expected values for a two-way table are calculated based on the assumption that Ho is true.

**♦ I can calculate the expected counts from a two-way table ♦**

The expected value for any cell in a 2-way table is:

|  |
| --- |
| E (X) = $ \frac{(row total)(column total)}{table total}$ |

Example 2: Find the expected counts for the table. Place them next to the observed values (in parentheses):

|  |  |  |  |
| --- | --- | --- | --- |
|  | Success | Failure | Total |
| Desipramine | 14 (8) | 10 (16) | 24 |
| Lithium | 6 | 18 | 24 |
| Placebo | 4 | 20 | 24 |
| Total | 24 | 48 | 72 |

Expected for Success Desipramine: $\frac{(24)(24)}{72}$ = 8

Expected for Failure Desipramine = $\frac{(24)(48)}{72}$ = 16

Finish the finding the rest of the expected counts and put these in the table above

**♦ I can calculate the Chi-Square statistic from a two-way table ♦**

The formula for Chi-Square test statistic is the same as it was in the previous section:

|  |
| --- |
| $$X^{2}= \sum\_{}^{}\frac{(Observed-Expected)}{Expected}^{2}$$ |

Example 3: Find the Chi-Squared value for the two way table in example 2.

 $X^{2}$ = $\frac{(14-8)^{2}}{8}$ + $\frac{(10-16)^{2}}{16}$ + $\frac{(6-8)^{2}}{8}$ + $\frac{(18-16)^{2}}{16}$ + $\frac{(4-8)^{2}}{8}$ + $\frac{(20-16)^{2}}{16}$ = 10.5

This value is then used to determine the P-value (probability that Ho would still be true with these results).

|  |
| --- |
| In a two way table the degrees of freedom is = (number of rows – 1) (number of columns – 1) |

Example 4: Determine the degrees of freedom and find the $X^{2}$ value for this test:

 Degrees of freedom = (rows – 1) (columns – 1) = (3 – 1) (2 – 1) = (2)(1) = 2 degrees of freedom

 Use 2 df and $X^{2}$ of 10.5 to find the P-value in your Chi-Squared table:

 This tells us that with these results, the probability that Ho is still true is between 0.005 and 0.01.

 Since this value is so low, we would not believe the Ho (reject the Ho).

**♦ I can perform a Chi-Square test to determine association ♦**

The steps to a Chi-Squared test are still the same (PHCTAC).

P: State the problem (Chi-Squared tests do not have parameters and statistics to identify) State the significance level (α)

H: Ho: There is no difference (or association or relationship) between the categories (in context)

 Ha: There is a difference (or association or relationship) between the categories (in context)

C: Conditions for $X^{2}$ : Random

All expected values must be at least 1

80% of expected values must be greater than 5 (usually all of these are > 5)

Graph a segmented bar graph to show the category distributions

T: Find Chi-Squared value using formula: $X^{2}= \sum\_{}^{}\frac{(Observed-Expected)}{Expected}^{2}$

 Find degrees of freedom

 Find the P-value in the Chi-Squared table

A: Compare the P-value to α and make a decision (do not reject Ho or reject Ho)

C: State your decision/conclusion in context

Example 5: 12,986 people were studied to determine if there was a relationship between heart disease and anger. All subjects were free from heart disease at the beginning of the study. The findings are summarized in the table below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Low Anger | Moderate Anger | High Anger | Total |
| Heart Disease | 53 | 110 | 27 | 190 |
| No Heart Disease | 3057 | 4321 | 606 | 7984 |
| Total | 3110 | 4431 | 633 | 8174 |

Test to determine if there is a significant relationship between anger level and heart disease (follow all steps)

P:

Ho:

Ha:

Conditions (include segmented bar graph)

T:

A:

C: