**Nominal vs. Ordinal categorical variables**

Categorical Variables: • Data whose measurement scale is inherently categorical.

• Nominal – categories with no logical ordering •

Ordinal – categories with a logical order / only two ways to order the categories (binary IS ordinal)

By examining the distributions of categorical variables, you can do the following: 1. Determine the frequencies of data values 2. Recognize possible associations among variables

**Association of categorical variables**

An association exists between two categorical variables if the distribution of one variable changes when the level (or value) of the other variable changes.

• If there is no association, the distribution of the first variable is the same regardless of the level of the other variable.

**How much of a change is required to believe there actually is a difference?**

**Null Hypothesis** • There is no association. • The distribution of one variable does not change across levels of another variable. •

**Alternative Hypothesis** • There is an association. • The distribution of one variable changes across levels of another variable.

\ **Chi-square distribution**

The **Chi-Square test** comes from the ��-distribution. • Characteristics of the �"-distribution:

1. **Bounded Below By Zero 2. Right Skewed 3. One set of Degrees of Freedom**

**Pearson’s Chi-square**

The Pearson �" test works for comparing any two categorical variables.

D.F. = (# Rows – 1)(# Columns – 1)

**Likelihood Ratio Chi-square**

The Likelihood Ratio �" test works for comparing any two categorical variables.

D.F. = (# Rows – 1)(# Columns – 1)

**Assumptions of above 2 tests(**assumptions to tests of association )

Both of the above tests have a sample size requirement.

• The sample size requirement is 80% or more of the cells in the crosstabulation table need expected count larger than 5.

When we don’t meet the assumption we can use **the Fisher’s exact test** that calculates all possible permutations of data

• Both the Pearson and Likelihood Ratio Chi-Square tests can handle any type of categorical variable – either ordinal, nominal, or both.

• However, ordinal variables provide us extra information since the order of the categories actually matters compared to nominal.

• We can test for even more with ordinal variables against other ordinal variables – whether two ordinal variables have a linear relationship as compared to just a general one.

**Mantel-Haenzsel Chi-square**

The Mantel-Haenszel �" test works for comparing any two ordinal variables. (DF= 1)

Chi-Square Tests

• Determines whether an association exists • MAY NOT measure the strength of the association • Can compare when sample size similar • Can NOT compare when sample size different

Measures of association: Measures the strength of the association

Odds ratios, - (Only for 2x2 tables – binary vs. binary): An odds ratio indicates how much more likely, with respect to odds, a certain event occurs in one group relative to its occurrence in another group

The odds of an event occurring is NOT the same as the probability that an event occurs

Odds = p/(1-p)

odds vs probability -

Cramer’s V (Any size table): when you have more than two categories in one or both variables use Cramer

Bounded between 0 and 1 (-1 and 1 for 2x2 scenario) where closer to 0 the weaker the relationship.

Spearman’s correlation - (ordinal vs. ordinal)

able to compare ordinal variables with any number of categories)

strength of association between two ordinal variables. pearman correlation is a correlation on the ranks of the observations as compared to the actual values of the observations.