Statistical Methods

Categorical Data – Counts, Percents, Proportions

Basic Categorical Inference

One Proportion Test (3.1)

The population is not divided into groups. There are only two possible outcomes for each response. We are comparing our sample proportion to a specific value.

Also called single proportion and 1P.

Two Proportion Test (3.2)

There are two groups with responses having two possible outcomes. We're comparing the proportion of the two groups to each other.

Also called difference in the proportions and 2P.

Advanced Categorical Inference

Chi-Square Goodness of Fit Test (3.3)

This is an extension of the one proportion test where there is a single variable but now there are more than two possible outcomes. You may either have specific proportions for each outcome or you can say all the outcomes occur equally.

The values represent the counts for each choice in the response variable.

Chi-Square Test for Association – Cross Tabulation (3.4)

This is an extension of the two proportion test. There are two variables used for classifying the data and you have a contingency table or a cross tabulation. Use this when you have more than 2 possible outcomes for either classification variable.

Two variables are used for classification purposes and the values represent the observed counts in each combination of the variables.

This is also called a chi-square test for independent or chi-square test for homogeneity. The table that summarizes the data is sometimes called a cross tabulation, contingency table, or joint frequency table.

Numeric Data – Measurements, Means, Averages

Basic Numeric Inference

One Sample T Test (4.1)

There is only numeric variable and the values are not grouped in any way. We are comparing the mean of those numbers to a specific value.

This is also called one mean, single mean, or 1T.

Paired Samples T Test – Dependent Samples (4.2)

This has two numeric variables for a single group. The two values belong to the same item and are paired together. We form a new variable by subtracting the two original variables. Then we look at the average of the differences.

Also called dependent samples or the mean of the difference.

Two Sample T Test – Independent Samples (4.3)

There is one numeric variable but there are two different groups. We are comparing the means of the groups to each other.

Also called 2T, independent means, and difference of the means.

Advanced Numeric Inference

One-Way ANOVA (4.4)

This is an extension of the two sample T test. There is still one numeric variable, but now we are comparing the means of three or more independent groups with each other.

Two-Way ANOVA (not in text)

There is one numeric variable with two classification variables. We are comparing the means of each grouping variable as well as whether there is any interaction between the variables.

Correlation & Regression (Ch 5-6)

Both of these use paired data similar to the paired samples T test. The difference is that here we do not expect that they will be equal, just that they might be related. **Correlation** measures the strength of a linear relationship while **Regression** describes the linear relationship by finding the equation of the line. *Regression can fall under either Basic or Advanced inference, depending on the number of variables.*

Colors indicate the Minitab Stat menu location: Basic Statistics, ANOVA, Tables, or Regression. Parentheses indicate the section from *Introductory Statistics with Randomization and Simulation*.