**Stat 401 - lab 12**

Goals:

 How to fit 2 way ANOVA models

 How to analyze data from a randomized complete block design

 How to analyze proportions in groups

We will use the collard data discussed in lecture. This has two factors: size and clarify, each with 2 levels. The data are in collard.txt. This (like all .txt files) has spaces between columns. Load the data

**How to fit 2 way ANOVA models**

 Very little new needed. Most of the information is where to find useful results.

 Do make sure that all variable are red bar variables (so define groups).

 In the collard data, size and clarify contain words, so they are automatically read as red bar.

 If a variable contains numbers (e.g., groups are indicated by 1, 2, 3), you need to convert to red bar.

 Analyze / Fit Model,

 Add both factors to the Construct Model Effects box

Add the interaction by clicking both factor names (either both from the variable list or one from the variable list and one from the Model Effects box), then Cross.

Leave the emphasis as Effect Leverage. We don't care about the leverage plots, but the Least Squares Means table is very helpful.

Click Run.

Where to find useful results:

 Overall ANOVA F test (the one-way ANOVA F test): Under Whole Model, Analysis of Variance table

 Residual vs Predicted value plot: Under Whole Model

 Pooled error standard deviation: In Summary of Fit box under Whole Model,

 look for Root Mean Square Error

 F tests for each factor and the interaction: Open the Effect Tests box (click the grey triangle)

 these are the type III (partial) tests for each main effect and the interaction

 Marginal means for each factor (averaged over all levels of the other factor)

 In the Least Squares Mean table for each factor (look to the top right of the Whole Model results)

 Important: You want the values in the Least Squares Means.

 The values under Means are a different sort of average that we haven't discussed.

 The individual treatment means: In the Least Squares Means table for the interaction.

 Contrasts or differences among LSMeans (either for a main effect or individual treatment means)

 Click the red triangle by the name of the factor.

 You should recognize the options. We used them with one way ANOVA.

 LS Means Contrast: you specify the coefficients

 LS Means Student t: all pairwise differences, no adjustment

 LS Means Tukey HSD (when more than 2 groups): all pairwise differences, Tukey adjustment

**Analyzing data from a randomized complete block design:**

 Same as a 2 way ANOVA model, except do not include the interaction

**Analyzing proportions**:

We will use the Vitamin C data used in lecture. This is in vitc.txt. Load the data.

Contingency table data can be provided in two different formats:

 As a summary table. The data set includes the combinations of row and column variables and a count.

 In the vitc data, n is the number of individuals with that combination of traits.

 As individual observations. You could instead provide information about each individual.

 e.g., Placebo Yes then Placebo Yes then Placebo No, …

 JMP will count up the numbers in each group and construct the contingency table for you.

For a 2 way contingency table (two factors or one factor and one response):

 Fit Y by X (Not Fit Model)

 Put one factor into the Y box and one factor into the X box.

 If one variable is a response (e.g., cold for the vitc study), then put it in the Y box.

For summary data: (Very important)

 Find the Freq box in the model dialog. Put the variable with the counts of observations in Freq.

 The model dialog should look like this:



For individual observations: all you need is Y and X, no Freq.

Important: If the contingency table has a total of 4 observations, 1 in each category, you forgot to fill in the Freq box to specify the count variable. (all JMP sees is four individual observations)

Where to find useful results:

 Counts for each cell in the contingency table: first row of each cell in the contingency table box

 Percentages: JMP gives you three percentages:

 total: as a percent of the total number (so out of 818 for the vitc data)

 row: as a percent of the row total (so out of 411 for the placebo group and 407 for the vitc group)

 column: as a percent of the column total (so out of 181 No or 637 Yes)

 Chi-square test: under Tests / test, look for the Pearson line

 This does not use a continuity correction.

 The likelihood ratio line is a very similar test preferred in some fields. Rarely different from Pearson

 Fisher exact test (use for small counts): under Tests / test, look for the 2-tail line

 Difference of proportions: need to hand calculate from percents in the contingency table

 se or confidence interval for difference of proportions: JMP doesn't do this for you

 Odds for each group: need to hand calculate from percents in the contingency table

 Odds ratio: Need to request this

 click the red triangle at the top, by Contingency table analysis

 select Odds Ratio (middle of the list)

 A box opens at the bottom of the results with the odds ratio and a 95% confidence interval

 The odds ratio is calculated as the odds of column 1 in row 1 divided by the odds of column 1 in row 2.

For the Vit C data, the odds of No (column 1) are:

 76/335 = 0.23 for the placebo group (row 1) and

 105/302 = 0.35 for the vitc group (row 2).

The odds ratio is 0.23 / 0.35 = 0.65, as reported.

If you want the odds ratio "the other way", just calculate the reciprocal of the odds ratio and its ci.

My advice is to look at the proportions and figure out which odds ratio you want (> 1 or < 1).