**To Copy and paste your data into JMP:**

Open your data in Excel

Place the data you want (not the entire spreadsheet) into your Excel copy buffer

invoke JMP separately from Excel using Start menu.

File-> New -> Data Table (^N)

once the new table appears, click Edit -> “Paste with Column Names”

To name the table, left click on the Table name in the left hand corner of the Table

Make sure JMP categorizes each variable correctly as continuous (blue ramp) (look for a blue ramp under Columns on the left hand navigation bar). The three data types JMP recognizes are continuous (blue ramp), ordinal, and nominal (categorical – red histogram).

**Descriptive (Summary) Statistics:**

Choose Tables > Summary > to open dialog box.

Click on the variable you want to group by (if any, e.g., “Species”); then click “Group”;

Click (highlight) on the variables for which you want the summary statistics (e.g., “DBH”)

and click on the summary statistics you want JMP to calculate (JMP won’t calculate the mode).

As you click, the stats you want will apper in the Statistics box

Click ‘OK’

**Correlation Analysis**

Click “Analyze” and go to “Multivariate Methods” options. Choose “Multivariate.”

Under the Multivariate pull-down menu choose “Pairwise comparisons” to get the p-value for the correlation.

**Simple Linear Regression.**

Using the Fit Y by X dialog box, enter your dependent and independent variables.

Using the red arrow to the left of the Bivariate fit box, click on “Fit Line” to fit a line through your cloud of points.

To make a **scatterplot** in JMP, alter the graph from the Fit Y by X output. Change the color of the regression line, alter the axes (right click on the graph to see the options for altering the graph). To annotate the figure to add the R2 and p-values: Remember there is a hidden toolbar at the top of each JMP graph/output. Click in the small horizontal space below the windows bar to see it. Once there, use the T (annotate) tool.

**t-test**

Analyze - Fit Y-by-X

Oneway Analaysis of Y (e.g., Height) by X (e.g., Site ID)

Click the Red arrow in box at top when graph appears, click t Test

Click Means/Anova/Pooled t , you will see both Oneway Anova and t-test.

Look for F Ration and Prob > F under Anova

**ANOVA**

Make sure that your independent variable (X) is categorical (red histogram), and that your dependent variable is continuous ( blue ramp). You can switch how JMP interprets data types by clicking on the red histogram (or blue ramp).

Note that ANOVA has two important assumptions: normality of within groups and equal variances among treatments.

Shapiro-Wilk test for normality

To check if your data fit the assumptions of ANOVA, run a Shapiro-Wilks test for normality for each treatment. This will take one separate test for each individual treatment because you will need to exclude all other values while testing for the normality of each individual treatment.

Exclude all but one treatment (e.g., Site) by highlighting the treatments to be excluded  
do this by clicking in the 1st column (the one that numbers rows), of the data table for all rows you want to exclude

then right clicking in that column and choosing “exclude” (you can also click on the red arrow in the ‘rows’ and then choose “exclude” from the drop down menu). Include/exclude is a ‘toggle’ switch, so to “include” after “excluding”, just click again!

then in JMP choose (Analyze->Distribution) for the variable of interest (e.g., Height ….) click OK

Once the histogram appears, use the red drop down arrow (for the variable of interest “…”), choose “Continuous fit” Normal? and

under the lowest drop down arrow (Fitted Normal)

choose “Goodness of Fit” to get the Shapiro Wilks test results.

Shapiro-Wilk tests for whether the data in your sample differ significantly from a normal distribution. If p is less than 0.05 for any of the treatments in your dataset, then your sample was likely not taken from a normal population and you want to fail to reject the null hypothesis. In this case, you should not proceed on to the ANOVA.

tired of scrolling to highlight all but one group? try using the data filter!!!

on rows menu let’s say you want to exclude Dan 1 genotype.

-> exclude all

->data filter

select genotype, then click “Add”

select Dan 1

ask me to show you! You will find that pop up menu in the same menu where you find Include/Exclude.

Levene’s test for homogeneity of variance

To run a Levene’s test for homogeneity of variance. You can find this Test for Unequal Variances after you run a “Fit Y by X” using your categorical variable as X and your continuous variable as Y (again, click on the red arrow above the graph window – look for Unequal Variances). If p is less than 0.05 it means your data are not equally variable, and you will want to fail Levene’s test.

I found it easiest to check first for equal variances – since if they are not equal, then I don’t bother to check for normality (which requires one per group!).

One-way ANOVA

To test the null hypothesis that all treatments are equal: using the “Fit y by x” function, place independent variable (aka “treatment”, here - number of species) as your X variable and dependent variable (aka “response” here – biomass) as Y to run an ANOVA – Under the red drop down arrow, choose “Means/ANOVA.” Report the F Ratio and p value (Prob > F). If p<.05 you have demonstrated among group variation dominates over within group variation, and can reject the null hypothesis.

Tukey’s HSD comparison of means

To determine which mean values are different from one another (if any).

On the Fit y by x graph,

click on the topmost red down arrow, and select “compare means”,

then All pairs, Tukey’s HSD.

(Cool intersecting (or not) circles will appear to the right of the graph!, with the label:

All Pairs Tukey-Kramer 0.05)

To interpret the test, look at the bottom, and you will see how each group compares with the other groups (convince yourself that all groups are compared!

‘Connecting Letters Report’ and ‘Ordered Differences Report’ tell the story!

Report F(x,x) = X.XX, p = X.XXXX and Tukey’s lowercase letters on the graph with the HISTOGRAM from JMP.

2-WAY ANOVA

Now, this isn’t the whole picture. Technically we have also treated half the trees with fertilizer. We want to know what the effect of genotype is, what the effect of fertilizer is, and if the effect of fertilizer is the same for all genotypes (which is analogous to asking if there a genotype x fertilizer interaction). For the purposes of this lab, let’s assume our data meet all ANOVA assumptions.

Using Fit Model , input decomposition rate as your Y variable: highlight both Genotype and Fertilization on the left and either click “Macros - full factorial” or paste both Genotype and Fertilization into X, then highlight both and click “Cross.” You ultimately want Genotype, Fertilization, and Genotype\*Fertilization in the “Construct Model Effects” box. Once you have that, click “run”.

You will want to report results such as those below:

Report your ANOVA results below:

**Summary of Fit**

|  |  |
| --- | --- |
| RSquare | .620145 |

**Analysis of Variance**

| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Ratio** |
| --- | --- | --- | --- | --- |
| Model | 9 | 4.48487e-6 | 4.9832e-7 | 6.1675 |
| Error | 34 | 2.7471e-6 | 8.0797e-8 | **Prob > F** |
| C. Total | 7.23197e-6 | 7.23197e-6 | XXXXXXXXXXX | .0001 |

**Effect Tests**

| **Source** | **Nparm** | **DF** | **Sum of Squares** | **F Ratio** | **Prob > F** |
| --- | --- | --- | --- | --- | --- |
| Genotype | 4 | 4 | 3.04875e-6 | 9.4333 | <.0001 |
| Fertilization | 1 | 1 | 5.1231e-7 | 6.3407 | 0.0167 |
| Genotype\*Fertilization | 4 | 4 | 8.80903e-7 | 2.7257 | 0.0453 |

While your JMP results output window is still open, go up to the red triangle and click on “Factor Profiling” then “Interaction plots.” This will provide a graphical output of the interactions in a window labeled “Interaction Profiles”. Expand the box so you can better see the patterns. Which lines in the box with labeled genotypes appear to cross one another?  Which genotypes appear unaffected by fertilization treatment?

**Report** the interaction profiles below.

Under “Chart” make a graph that includes mean decomposition rate (as Y) and both Genotype and Fertilization as (X) categories. Paste the graph below (be sure to include error bars). To make decomp rate as Y, you will need to highlight it, then pull down “Statistics” and click Mean.

**Ask yourself if** the patterns shown in the bar chart above matches the interaction plots from Q9?

3-WAY ANOVA

As with 1-way and 2-way ANOVAs, be sure that JMP interprets the data type of all independent (“x”) variables as nominal (or categorical ). In addition, as in all parametric ANOVAs, you will need to assure that the independent variables are normal and that variances are equal.

An example: to see the effects of Genotype and Fertilization change through time (making time a categorical variable, even though it is usually considered continuous). Look at % mass remaining at each harvest date to look for these patterns. Using the “ln-transformed mass remaining” data as Y, test for the effects of Genotype, Fertilization, and Harvest Date (categorical).

This test requires that we also examine all possible interactions among our variables: G x F, G x D, F x D. Using “Fit Model”, input all these variables (hint - highlight all three categorical variables on the left and use the Macro - full factorial).

**For Chi Square (with 1 variable – expected vs. not, as with fish in substrate example)**

Paste your data into JMP as usual, taking care to reorganize the data into the format required by JMP (see below). Choose “Analyze” – “Distribution,”(putting your variable into Y) then ask JMP to create a mosaic plot for you (for this example, the variable was “substrate”, so you will see mosaic plot under the “substrate” tab ). Also ask JMP to “test probabilities” (also under the substrate tab) – this is where you can enter the expected probabilities of numbers of fish seen in each habitat type (hint – if 50% of the substrate is sand, we would expect to find 50% of the fish in sandy habitats – if the null hypothesis is true – make sure you type the correct percentage in the correct cell!).

report the Chi-square, df, p, and “interpretation”, e.g.,

**For Chi Square. (with 2 variables – as with the cactus example)**

Using the “Fit Y by X” function, enter cactus as your Y variable and shrub as your X variable (you could easily switch these and you would come up with the same answer in this example).

**How to organize your data using Excel so it can be used in a Chi-Square:**

typically, one sees chi-square data organized as a table (see below)

|  |  |  |
| --- | --- | --- |
|  | # fish observed | # fish expected |
| sand (50%) | 8 | 15 |
| gravel (30%) | 18 | 9 |
| silt (20%) | 4 | 6 |

However, for JMP, you will need to organize such data as observations, e.g.,

substrate (header for new table – for this example, one column)

sand (repeat this row 8 times)

|  |
| --- |
| substrate |
| sand |
| sand |
| sand |
| sand |
| sand |
| sand |
| sand |
| sand |

gravel (repeat 18 times)

|  |
| --- |
| gravel |
| gravel |
| gravel |
| …. |

silt (repeat 4 times)

|  |
| --- |
| silt |
| silt |
| silt |
| …. |

for two way anova - (???)

|  |  |  |  |
| --- | --- | --- | --- |
|  | Shrub yes | Shrub no | row totals |
| Cactus yes | 3 | 7 | 10 |
| Cactus no | 5 | 88 | 93 |
| column totals | 8 | 95 | 103 |

|  |  |
| --- | --- |
| cactus | shrub |
| yes | yes |
| yes | no |
| no | yes |
| no | no |

first row repeated 3 times

2nd row – 7 times, etc. – whoops in ss it’s 6 times!!!

3rd – 5 times

4th – no-no 88 times!

When you have two columns, e.g., above, use the “Fit Y by X” function, enter one column as your Y variable and the other as your X variable (if one is likely to depend on the other, it should be the Y variable). As before you will calculate the chi square, but you also get additional information, namely which is preferred.

| **Fisher's Exact Test** | **Prob** | **Alternative Hypothesis** |
| --- | --- | --- |
| Left | 0.9973 | Prob(cactus=yes) is greater for shrub=no than yes |
| Right | 0.0290\* | Prob(cactus=yes) is greater for shrub=yes than no |
| 2-Tail | 0.0290\* | Prob(cactus=yes) is different across shrub |

**CART model, paste the data into JMP as usual**

Choose “Analyze – Modeling” then “Partition” modeling (JMP calls a CART model “Partition modeling”). Choose the dependent variable as your Y variable and enter all the continuous trait variables as ‘X’s. Then, ask JMP to “split” your tree – to make the first dichotomous divide.

Keep ‘splitting’ until you are happy with your R-square; there are no hard and fast rules. If you specify “Fit details”, you will see information about how good the fit is.

LOGISTIC REGRESSION

for logistic regression you would have a categorical Y variable (just like all other cases today), but you would have a continuous X variable.

For example, in a study of rabbits exhibiting some type of infection, you might want to know what dose of penicillin results in a 0.5 probability of curing a rabbit (think of Chester Bliss’ probit analysis here…).

Using the data provided in the tab “Advanced-rabbits” import the data into JMP. Using the “Fit Y by X” platform, use “response” as Y and “ln(dose)” as X. JMP will automatically fill in the column “count” as a frequency – make sure this happened, and if not, place “Count” in the frequency box.

SOME USEFUL JMP TRICKS:

1. Name your tables (so you can just keep them around, or save them).
2. Figure out how to stop a session, and save it, so you can resume later, perhaps on a different computer.