Chp 11
ANOVA

Example:
Weights (kg) of Poplar Trees

<table>
<thead>
<tr>
<th>Treatment: None</th>
<th>Fertilizer</th>
<th>Irrigation</th>
<th>F and I</th>
<th>k=4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Factor)</td>
<td>0.15</td>
<td>0.02</td>
<td>0.46</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>0.02</td>
<td>0.14</td>
<td>0.02</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>0.16</td>
<td>0.22</td>
<td>0.08</td>
<td>0.05</td>
</tr>
<tr>
<td>(M1)</td>
<td>(M2)</td>
<td>(M3)</td>
<td>(M4)</td>
<td></td>
</tr>
</tbody>
</table>

Q: whether $M_1 = M_2 = M_3 = M_4$

If we want to test $M_1 = M_2 = M_3 = M_4$, and we use two sample test method,
we are going to test the following b. Hypothesis:
Ho: $M_1 = M_2, M_3 = M_4$
Ha: $M_1 = M_4, M_2 = M_3$

\[
\begin{array}{c|c|c}
\text{Ho is true} & \text{Not Reject} & \text{Reject} \\
\hline
.95 & .05 \\
\text{Not true} & \text{Typ II Error} & \checkmark
\end{array}
\]

\[
\begin{align*}
P(\text{Not Reject } \text{Ho}_1) &= .95 \\
P(\text{Not Reject } \text{Ho}_2) &= .95 \\
P(\text{Not Reject all 6 Ho's}) &= .95^6 \\
&= .735 \\
P(\text{H}_{06}) &= .95
\end{align*}
\]
\( X \) denotes the \# of Not Reject the Hypothesis (two sample),
\[ X \sim \text{Bin}(6, 0.95) \]

\[(0.95)^6 = 0.735\]
\[(?)^6 = 0.95\]
\[(0.9915)^6 = 0.95\]

Interactive:

Weights

\[
\begin{align*}
\text{Factor}_1 & \quad \text{Factor}_2 \\
0 & \quad 1 \\
\text{(No Fertilizer)} & \text{(Fertilizer)}
\end{align*}
\]

\( \text{red} X \): Irrigation 1

\( \leftrightarrow \) No interactive

Interactive
One-way ANOVA

treatment (factor): a property, or characteristic, that allows us to distinguish the different populations from one another.

Balanced Design: Same # of samples for each group (sub group)

Read the Output.

F-distribution.  
\[ d.f. \ (k-1, \ N-k) \]

F

k is the # of groups.
N is the total # of samples.

Two-way ANOVA

Read the Output.
1. Do parents know their teenagers' drug use? When 130 parents were asked if they thought that their teenager had tried marijuana, 28 said yes. When 156 teenagers were asked if they had tried marijuana, 58 said yes. At the 0.05 level, test the claim that the proportion of parents who think their teenager has tried marijuana is equal to the proportion of teenagers who say that they have tried marijuana. Note that the formula for the test statistic is \[ Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \] where \( \hat{p} = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2} \).

\[ H_0: \hat{p}_1 = \hat{p}_2 \]
\[ H_a: \hat{p}_1 \neq \hat{p}_2 \]

\[ Z = \frac{\frac{28}{130} - \frac{58}{156}}{\sqrt{\frac{\hat{p}(1-\hat{p})}{130} + \frac{\hat{p}(1-\hat{p})}{156}}} \]

\[ \hat{p} = \frac{28/130 + 58/156}{130 + 156} = 0.3 \]

\[ Z = \frac{0.218 - 0.386}{\sqrt{0.3(1-0.3)(1/130 + 1/156)}} = -2.874 \]

March 8: \( Z = \frac{-2.874}{\sqrt{2847}} = -2.89 \)

Comparing with 1.96, \( Z \) is less than the critical value for a two-tailed test at the 0.05 level. Therefore, we reject the null hypothesis.

The proportion of parents is NOT equal to the proportion of teenagers who say that they have tried marijuana.

Total: 15 points
Measuring the height of tall redwood trees can be difficult. One way to get a good estimate is to use the fact that the height (in feet) of a redwood tree can be related to the diameter (in inches) of the tree at five feet off of the ground. Here is JMP output from a regression of tree height on tree diameter, with questions on the back of this page.
1. State the hypotheses for testing for a significant relationship between height and diameter. Be sure to define your notation.

\[ H_0: \beta_i = 0 \quad \text{\(\beta_i\) is the slope or coefficient} \]

\[ H_a: \beta_i \neq 0 \]

\[ H_0: \rho = 0 \quad \text{\(\rho\) is the correlation coefficient} \]

\[ H_a: \rho \neq 0 \]

2. What do you conclude from this hypothesis test? State your conclusion in the context of the problem.

From the output,

\[ r = \sqrt{0.728799} \]

\[ t = \sqrt{\frac{1-r^2}{n-2}} \]

\[ \text{d.f.} = n - 2 = 21 - 2 = 19 \]

or using table A

\[ \text{Reject } H_0: \beta_i = 0 \]

Hence there's a linear relation.

3. What is the fitted regression line?

\[ \text{Height} = 78.796 + 2.67 \times \text{Diameter} \]

4. What is the predicted height of a tree that has a diameter of 24 inches?

\[ \text{Height} = 78.796 + 2.67 \times 24 \]

5. Interpret the value of the fitted slope.

the height increases 2.67 feet as the diameter increases per inch.

6. Interpret the value of the \(R^2\). How good is this fit?

\[ R^2 = 0.728799 \]

The percentage of the total variation that has been explained by the linear regression.

7. How does the plot of the residuals look? It is pretty good.

Looks good, no particular pattern.