Overview

- Let’s continue our journey through the ANOVA approach to data
- Focus on Single Factor Models
- Terms for the ANOVA Table
- **R-square**
- More single factor models
- Strategies for Multiple Comparisons, including Fisher’s LSD

What are the Sum of Squares called?

<table>
<thead>
<tr>
<th>Terms Explained</th>
<th>Excel</th>
<th>JMP</th>
<th>SAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>SST - Sum of Squares Treatment</td>
<td>Between</td>
<td>Variable Name</td>
<td>Model</td>
</tr>
<tr>
<td>SSE - Sum of Squares Error</td>
<td>Within</td>
<td>Error</td>
<td>Error</td>
</tr>
<tr>
<td>SSTotal - Total Sum of Squares</td>
<td>Total</td>
<td>Total</td>
<td>Corrected Total</td>
</tr>
<tr>
<td>Factor Levels</td>
<td>Groups</td>
<td>Factors</td>
<td>Class</td>
</tr>
</tbody>
</table>
R-square

- The baseline model is the Grand Mean
  \[ SS_{Total} = \sum_{i=1}^{n} (y_i - \bar{Y})^2 \]

- Our model is one that is based on knowledge of the Factors/treatments
  \[ SST = \sum_{i=1}^{k} n_i (\bar{y}_i - \bar{Y})^2 \]

- R2 is a measure of the percent of the SS\(_{Total}\) that is due to the treatment
  \[ R^2 = \frac{SST}{SS_{Total}} \]

Another Problem

- An experiment is conducted to determine the differences in mean increases in plant growth from 5 different inoculums
- Inoculums are substances injected into a plant to fight disease.
- The experiment involved 20 cuttings of a shrub (all of equal weight), with 4 cuttings assigned to the five different inoculums
- The data represent the increase in weight in grams
- We will use \(\alpha = .05\)

Incoculum Data

- Here is the way Excel would prefer the data
  \[
  \begin{array}{c|c|c|c|c|c}
  \hline
  & I1 & I2 & I3 & I4 & I5 \\
  \hline
  15 & 21 & 22 & 10 & 6 \\
  18 & 13 & 19 & 14 & 11 \\
  9 & 20 & 24 & 21 & 15 \\
  16 & 17 & 21 & 13 & 8 \\
  \hline
  \end{array}
  \]
- We can add the means and variances
- And a box plot

R-square

- With \(R^2\) we ask, "how much better do I understand the Response variable (dependent) by knowing something about the Factors/Treatments (independent variables)
- \(R^2\) varies from 0 to 1
  - 0 means we explain nothing of the dependent variable
  - 1 means we explain it perfectly
- \(R^2\) is a linear measure of association
- \(R^2 = \)
  - \(\frac{SST}{SS_{Total}}\), or
  - \(1 - \frac{SSE}{SS_{Total}}\)
Excel results

- The results show that F* is 5.285 which has a p-value of .007
- Excel does not give us R2, but it is easy to calculate:
  \[ R^2 = \frac{292.80}{500.55} = .58496 \]
- 58.5% of the variability in GROWTH is due to the type of inoculum
- I can also solve R2 as
  \[ 1 - \frac{207.75}{500.55} = .58496 \]

ANOVA Hypothesis Test for Incoculm Data

- Ho: \[ \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 \]
- Ha: \[ \text{At least two means are different} \]
- Assumptions
  Equal variances, normal distribution
- Test Statistic
  \[ F* = 5.285 \quad p = .007 \]
- Rejection Region
  \[ F_{.05, 4, 15} = 3.056 \]
- Conclusion:
  F* > \[ F_{.05, 4, 15} \]
  or p = .007
  Reject Ho: \[ \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 \]

What’s Next? Compare which means are different

- ANOVA just tests that at least two of the means are different
- ANOVA does not tell us which means are different
- The next logical step is to ask which means are different from each other
- We have five levels of the factor
- Resulting in 10 different comparisons of treatment means
  - 1 to 2; 1 to 3; 1 to 4; 1 to 5;
  - 2 to 3; 2 to 4; 2 to 5;
  - 3 to 4; 3 to 5;
  - 4 to 5
### Difference of Means with Multiple Comparisons

- When we conduct a hypotheses test from a single experiment or sample, we set a level of Type I error for a comparison of two means.
- However, when we make many comparisons across treatments, the level of alpha increases in response to the number of comparisons.
- This is referred to as Experiment-Wise Error Rate (aka, family-wise error rate).
- \( \alpha = 1-(1-\alpha)^c \)
- where \( e \) is the experiment-wise error rate and \( c \) is the number of independent comparisons.

### Experiment-Wise Error Rate with Multiple Contrasts

<table>
<thead>
<tr>
<th>Number of Contrasts</th>
<th>Probability of a Type I Error on an Individual Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.10 0.05 0.01</td>
</tr>
<tr>
<td>2</td>
<td>0.19 0.09 0.02</td>
</tr>
<tr>
<td>3</td>
<td>0.27 0.14 0.03</td>
</tr>
<tr>
<td>4</td>
<td>0.34 0.18 0.03</td>
</tr>
<tr>
<td>5</td>
<td>0.41 0.20 0.04</td>
</tr>
<tr>
<td>6</td>
<td>0.49 0.26 0.05</td>
</tr>
<tr>
<td>7</td>
<td>0.52 0.20 0.06</td>
</tr>
<tr>
<td>8</td>
<td>0.57 0.33 0.07</td>
</tr>
<tr>
<td>9</td>
<td>0.61 0.37 0.08</td>
</tr>
<tr>
<td>10</td>
<td>0.65 0.40 0.09</td>
</tr>
<tr>
<td>11</td>
<td>0.68 0.43 0.10</td>
</tr>
<tr>
<td>12</td>
<td>0.71 0.46 0.11</td>
</tr>
</tbody>
</table>

### Fisher’s Least Significant Difference (LSD)

- Fisher developed a strategy to deal with this issue using the concept of the Least Significant Difference (LSD).
- In this approach, an alpha rate is fixed and a least significant difference is calculated.
- Fisher’s strategy was to develop a difference from which each comparison can be compared.
- The difference between two means would need to be at least the size of the LSD when the overall level of alpha would be fixed at the desired level.

### Fisher’s LSD

\[
LSD_{ij} = t_{a/2} \sqrt{\frac{s^2}{n_i} + \frac{s^2}{n_j}}
\]

- \( \alpha \), the desired level of Type I error for each comparison. This level of \( \alpha \) is fixed by the LSD approach.
- \( t_{a/2} \) a t-value associated with degrees of freedom error in the ANOVA table (set for a two tailed test in this example).
- \( s^2 \) the estimate of the pooled variance (MSE) from the ANOVA Table.
- \( n_i \) the sample size for group i.
- \( n_j \) the sample size for group j.
- In the case where the sample size is the same for each group, we calculate a single LSD using

### Fisher’s LSD for Inoculum Data

- \( \alpha = .05 \)
- \( t_{a/2}, 15 \text{ d.f.} = 2.131 \)
- \( MSE = 13.85 \)
- \( n \) for all groups = 4

\[
LSD = 2.131 \sqrt{\frac{13.85}{4}} = 5.6078
\]

- Order means from lowest to highest:

  - Order means from lowest to highest:
  - INC5 to INC3 21.50 - 14.50 > LSD
  - INC5 to INC1 21.50 - 14.00 > INC3
  - INC5 to INC2 17.75 - 10.00 > INC4
  - INC5 to INC4 14.5 - 10.0 = 4.5 < LSD
  - INC1 to INC3 21.50 - 14.50 > INC3
  - INC4 to INC3 21.50 - 14.50 > INC3
  - INC1 to INC2 17.75 - 14.5 = 3.25 < LSD
  - INC2 to INC3 21.50 - 17.75 = 3.75 < LSD

INC5 has the highest mean at 21.50.
INC3 is significantly different from INC5, INC1, and INC4.
INC2 is significantly different from INC5.
No other means were significantly different from each other.
All comparisons were significant at \( \alpha=.05 \) controlling for multiple comparisons using Fisher’s LSD.
Look at how JMP does this test

- The first matrix shows the difference minus the LSD
- Values that are positive show a difference that is significant
- The second table is also a popular way to show the same results
- Move down the columns to find significant differences

Experiment-Wise Error Rate

- There are many other methods of comparison
  - Scheffe
  - Tukey
  - Bonferroni
  - Tukey-Kramer
- Most of the multiple comparison strategies use the following approach
  1. Fix alpha at some level
  2. Adjust the comparisons to reflect an overall alpha
  3. Compare the selected means (or all of them) using a difference of means test using a pooled variance
  4. Many show the result in terms of a confidence interval
  5. If the Confidence Interval overlaps with zero – there isn’t a difference

Example for you to run and work on - ANOVA Golf.xls or ANOVA Golf.jmp

- The USGA wants to compare the mean distances of several brands of golf balls struck by a driver.
- They set up an experiment where a 10 balls are randomly picked from an allotment of four different brands of golf balls.
- To hold constant the effect of the golfer, they use a mechanical robotic golfer using the same driver.
- The distance the ball traveled is recorded as the response variable.
- Use an alpha level of .01.

Experimental Design

| 1 Factor: Golfball Brand |
| 4 Treatments |
| 10 replications per treatment |

Experimental Unit: Golfball
Measurement Unit: Golfball
Total Sample Size: 40

Results from Excel

- This is the way Excel prefers the data
- Looking at the means, I see Ball C went the furthest on average, and Ball D the shortest
- The Variances are similar - no ratio greater than 2.2
- I used TOOLS, DATA ANALYSIS, ANOVA Single Factor to run the ANOVA

<table>
<thead>
<tr>
<th>Ball A</th>
<th>Ball B</th>
<th>Ball C</th>
<th>Ball D</th>
</tr>
</thead>
<tbody>
<tr>
<td>251.2</td>
<td>263.2</td>
<td>269.7</td>
<td>251.6</td>
</tr>
<tr>
<td>245.1</td>
<td>262.9</td>
<td>263.2</td>
<td>248.6</td>
</tr>
<tr>
<td>248.0</td>
<td>260.0</td>
<td>277.5</td>
<td>249.4</td>
</tr>
<tr>
<td>251.1</td>
<td>254.5</td>
<td>267.4</td>
<td>242.0</td>
</tr>
<tr>
<td>260.5</td>
<td>264.3</td>
<td>270.5</td>
<td>246.5</td>
</tr>
<tr>
<td>250.0</td>
<td>257.0</td>
<td>265.5</td>
<td>251.3</td>
</tr>
<tr>
<td>263.9</td>
<td>262.8</td>
<td>270.7</td>
<td>261.8</td>
</tr>
<tr>
<td>244.6</td>
<td>264.4</td>
<td>272.9</td>
<td>249.0</td>
</tr>
<tr>
<td>254.6</td>
<td>260.6</td>
<td>275.6</td>
<td>247.1</td>
</tr>
<tr>
<td>248.8</td>
<td>255.9</td>
<td>296.5</td>
<td>245.9</td>
</tr>
</tbody>
</table>

| Mean | 251.00 | 261.63 | 270.33 | 249.70 |

- $F^* = 43.989, p < .001$
- $R^2 = 2794.39/3556.69 = .7857$
- 78.6% of the variability in driving distance is due to the ball type
**ANOVA Hypothesis Test for Golfball Data**

- **Ho:** $\mu_1 = \mu_2 = \mu_3 = \mu_4$
- **Ha:** At least two means are different
- **Assumptions:** Equal variances, normal distribution
- **Test Statistic:** $F^* = 43.989, p < .001$
- **Rejection Region:** $F > F_{.01, 3, 36}$ or $p < .001$
- **Conclusion:** Reject $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$

**Next: Which Golfballs are different from each other?**

- We will use Fisher’s LSD
  
  - There are 6 contrasts
  - $\alpha = .01$
  - MSE = 21.18
  - $n = 10$
  - $t_{.01/2, 36} = 2.719$

<table>
<thead>
<tr>
<th>Ball D</th>
<th>Ball A</th>
<th>Ball B</th>
<th>Ball C</th>
</tr>
</thead>
<tbody>
<tr>
<td>249.70</td>
<td>251.00</td>
<td>261.63</td>
<td>270.33</td>
</tr>
</tbody>
</table>

  - LSD = $2.719 \sqrt{\frac{2 \times 21.18}{10}} = 5.59612$
  - Ball C has the highest mean distance at 270.33.
  - Ball C is significantly different from Ball D, Ball A, and Ball B.
  - Ball B is significantly different from Ball A and Ball D.
  - No other means were significantly different from each other.
  - All comparisons were significant at $\alpha = .01$ controlling for multiple comparisons using Fisher’s LSD.

**Results from JMP**

- JMP (any advanced software) gives a complete analysis
- R-square
- ANOVA and $F^*$
- Mean Comparisons
- Software like JMP would also
- Test the assumption about equal variances
- Different Mean comparisons

**Summary**

- We looked at some more single-factor problems and the way to look at the results
- We introduced $R^2$ as a measure of association, which shows us how much of the variability in the response variable is explained by the factor levels.
- After we establish some of the treatment means differ from each other, we want to know which means are different.
- To do this we use a of “Experiment or Family-wise error rate” to make multiple comparisons of differences of means.
- We introduced Fisher’s LSD as a simple way to make multiple comparisons and control the overall “Experiment-Wise Error Rate.”