Chapter 16
One-way Analysis of Variance
PSY 295 – Oswald

Major Points
• An example
• The logic
• Calculations
  – Unequal sample sizes
• Multiple comparisons
  – Fisher’s LSD
  – Bonferron t

Major Points
• Assumptions of analysis of variance
• Magnitude of effect
  – eta squared
  – omega squared
  – d
• Review questions

Logic of the Analysis of Variance
• Null hypothesis $H_0$: Population means
  $\mu_1 = \mu_2 = \mu_3 = \mu_4$
• Alternative hypothesis: $H_1$
  –
  $\mu_1 > \mu_2 = \mu_3 = \mu_4$ OR $\mu_1 > \mu_2 = \mu_3 = \mu_4$
  OR $\mu_1 = \mu_2 = \mu_3 < \mu_4$ OR $\mu_1 > \mu_2 > \mu_3 = \mu_4$
  – etc.
Logic

• Create a measure of variability
  – MS\(_{\text{groups}}\)
• Create a measure of variability
  – MS\(_{\text{error}}\)

Logic

• Form ratio of \(\frac{MS_{\text{groups}}}{MS_{\text{error}}}\)
  – Ratio if null true
  – Ratio if null false
  – “approximately 1” can actually be as high as

Epinephrine and Memory

• Based on Introini-Collison & McGaugh (1986)
  – Trained mice to go on \(\text{Y}\) maze
  – Injected with 0, .1, .3, or 1.0 mg/kg epinephrine
  – Next day trained to go in same \(\text{Y}\) maze
  – dep. Var. =
    • indicates better retention of Day 1
    • Reflects epinephrine’s effect on memory

<table>
<thead>
<tr>
<th>Dosage ((\mu)g/kg)</th>
<th>0.0</th>
<th>0.1</th>
<th>0.3</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>7</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>6</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>8</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Mean

<table>
<thead>
<tr>
<th></th>
<th>0.0</th>
<th>0.1</th>
<th>0.3</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>3.22</td>
<td>4.50</td>
<td>5.50</td>
<td>1.89</td>
</tr>
<tr>
<td>St. Dev.</td>
<td>1.20</td>
<td>1.29</td>
<td>1.13</td>
<td>1.88</td>
</tr>
</tbody>
</table>

Grand mean = 3.78
Chapter 16  One-way Analysis of Variance

Calculations

- Start with Sum of Squares (SS)
  - We need:
    - $S_{\text{total}}$
    - $S_{\text{groups}}$
    - $S_{\text{error}}$
- Compute degrees of freedom ($df$)
- Compute

Summary Table

<table>
<thead>
<tr>
<th>Source</th>
<th>$df$</th>
<th>$SS$</th>
<th>$MS$</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groups</td>
<td>3</td>
<td>132.556</td>
<td>44.185</td>
<td>35.816*</td>
</tr>
<tr>
<td>Error</td>
<td>68</td>
<td>83.889</td>
<td>1.234</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>71</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* $p < .05$
Conclusions

• The $F$ for groups
  − We would obtain an $F$ of this size, less than one time out of 1000 (.001).
  − The difference in group means
  − The number of trials to learn reversal

Conclusions

• The injection of epinephrine following learning appears to that learning.
• High doses may have effect.

Unequal Sample Sizes

• With one-way, no particular problem
  − by appropriate $n_i$ as you go
  − The problem is more complex with more complex designs.
• Example from Foa, Rothbaum, Riggs, & Murdock (1991)

Post-Traumatic Stress Disorder

• Four treatment groups given psychotherapy
  ◆ Stress Inoculation Therapy (SIT)
  ◆ Standard techniques for handling stress
  ◆ Prolonged exposure (PE)
  ◆ Reviewed the event repeatedly in their mind
  ◆ Supportive counseling (SC)
  ◆ Standard counseling
  ◆ Waiting List Control (WL)
  ◆ No treatment
Chapter 16  One-way Analysis of Variance

### Tentative Conclusions

- Fewer symptoms with SIT and PE than with other two
- Also within treatment groups
- Is just a reflection of variability of individuals?

### Calculations

- Almost the same as earlier
  - Note differences
- We multiply by $n_j$ as we go along.
- $MS_{\text{error}}$ is now a .

<table>
<thead>
<tr>
<th>SIT</th>
<th>PE</th>
<th>SC</th>
<th>WL</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>13</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>21</td>
<td>34</td>
<td>11</td>
</tr>
<tr>
<td>14</td>
<td>21</td>
<td>5</td>
<td>17</td>
</tr>
<tr>
<td>30</td>
<td>27</td>
<td>20</td>
<td>27</td>
</tr>
<tr>
<td>17</td>
<td>17</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>8</td>
<td>17</td>
<td>34</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>17</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>12</td>
<td>23</td>
<td>19</td>
<td>13</td>
</tr>
<tr>
<td>13</td>
<td>21</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>15</td>
<td>27</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>18</td>
<td>25</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>23</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>19</td>
<td>19</td>
<td>13</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>13</td>
<td>8</td>
</tr>
</tbody>
</table>

Mean = 11.071  15.400  18.091  19.500

St. Dev.  3.951  11.118  7.134  7.106

Grand mean = 15.622

### Calculations--cont.

\[
SS_{\text{total}} = \Sigma (X - \overline{X})^2 = (3-15.62)^2 + (13-15.62)^2 + \ldots + (13-15.62)^2
\]
\[
= 2786.6
\]

\[
SS_{\text{groups}} = \Sigma n_j (\overline{X}_j - \overline{X})^2
\]
\[
= \left[ 14(11.071-15.62)^2 + 10(11.118-15.62)^2 + \ldots + 10(19.500-15.62)^2 \right]
\]
\[
= 507.8
\]

\[
SS_{\text{error}} = SS_{\text{total}} - SS_{\text{groups}} = 2786.6 - 507.8
\]
\[
= 2278.8
\]
Summary Table

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groups</td>
<td>3</td>
<td>507.8</td>
<td>169.3</td>
<td>3.05*</td>
</tr>
<tr>
<td>Error</td>
<td>41</td>
<td>2278.7</td>
<td>55.3</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>44</td>
<td>2786.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>* p &lt; .05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ F_{0.05}(3,41) = 2.84 \]

Conclusions

- \( F \) is significant
- The population means
- In general
  - \( SIT \) most effective.

Multiple Comparisons

- Significant \( F \) only shows that not all groups are equal
  - We want to know which are different.
- Such procedures are designed to control familywise error rate.
  - defined
  - Contrast with error rate

More on Error Rates

- Most tests use a significance level (\( \alpha \)) for each \( t \) test.
- The more likely we are to make Type I error.
  - Good reason to
Fisher’s LSD Procedure

- Requires significant overall $F$, or no tests
- Run standard $t$ tests between pairs of groups.
  - Often we replace $s^2_j$, or pooled estimate with $\text{MS}_{\text{error}}$ from overall analysis
  - It is really just a pooled error term, but with -pooled across all treatment groups.

Bonferroni $t$-test

- Run $t$ tests between , as usual
  - Hold down number of $t$ tests
  - Reject if $t$ exceeds critical value in Bonferroni table
- Works by using a more for each comparison

Bonferroni $t$-test

- Critical value of $\alpha$ for each test set at , where $c = \text{number of tests run}$
  - Assuming familywise
  - e.g. with 3 tests, each $t$ must be significant at level.
- Make sure calculated probability <
- Necessary table is in the book

Assumptions for ANOVA

- Assume:
  - Observations within each population
  - Population variances are of variance or homoscedasticity
  - Observations are
Assumptions

• Analysis of variance is generally
  – A robust test is one that is not greatly affected by
    of assumptions.

Magnitude of Effect

• Eta squared ($\eta^2$)
  – Easy to calculate
  – on the high side
  – Formula follows...
  – Percent of variation in the data that can be attributed to

$\eta^2 = \frac{SS_{\text{groups}}}{SS_{\text{total}}} = \frac{507.8}{2786.6} = .18$

$\omega^2 = \frac{SS_{\text{groups}} - (k-1)MS_{\text{error}}}{SS_{\text{total}} + MS_{\text{error}}} = \frac{507.8 - 3(55.6)}{2786.6 + 55.6} = .12$

• $\eta^2 = .18$: 18% of variability in symptoms can be accounted for by treatment
• $\omega^2 = .12$: This is a less biased estimate, and note that it is 33% smaller.

Magnitude of Effect

• Omega squared ($\omega^2$)
  – Much than $\eta^2$
  – Not as intuitive
  – We both numerator and denominator with $MS_{\text{error}}$
  – Formula on next slide
Other Measures of Effect Size

- We can use the same kinds of measures we talked about with $t$ tests.
- Usually makes most sense to talk about $\eta^2$, rather than a measure averaged over several groups.

Review Questions

(For review – these are NOT the only things to know)

- Why is it called the analysis of “variance” and not the analysis of “means”?
- What do we compare to create our test?
- Why would large values of $F$ lead to rejection of $H_0$?
- What do we do differently with unequal $n$'s?

- What is the per comparison error rate?
- Why would the familywise error rate generally be larger than .05 unless it is controlled?
- Most instructors hate Fisher’s LSD. Can you guess why?
  – Why should they not hate it?

- How does the Bonferroni test work?
- What assumptions does the analysis of variance require?
- What does “robust test” mean?
- What do we mean by “magnitude of effect”?
- What do you know if $\eta^2$ is .60?
Review Questions
(For review – these are NOT the only things to know)

• Why would we calculate effect size measures on only pairs of groups at a time?