Analysis of Variance (ANOVA)

- Treatment effect
  - The existence of at least one difference between the population means categorized by the independent variable
- Random error
  - The combined effects of all uncontrolled factors (on the scores of subjects)

Analysis of Variance (ANOVA)

- $F$ ratio:
  \[ F = \frac{\text{Variability between groups}}{\text{Variability within groups}} \]

Analysis of Variance (ANOVA)

- If the $H_0$ really is true
  \[ F = \frac{\text{random error}}{\text{random error}} \]
- If the $H_0$ really is false
  \[ F = \frac{\text{treatment effect} + \text{random error}}{\text{random error}} \]

Analysis of Variance (ANOVA)

- $F$ test:
  - It is based on the notion that if the null hypothesis really is true, both the numerator and the denominator of the $F$ ratio will tend to be similar
  - But if the null hypothesis really is false, the numerator will tend to be larger than the denominator

Analysis of Variance (ANOVA)

- This whole ANOVA business looks pretty complicated and tedious …
- If we want to analyze more than two population means, couldn’t we simply perform several $t$ tests comparing pairs of population means?
Analysis of Variance (ANOVA)

- The answer is NO, because that would increase the Type I error rate:

\[ 1 - (1 - \alpha)^c \]

- \( \alpha \) = level of significance of each separate test
- \( c \) = number of independent t tests

Therefore, we have to use a test that deals with more than two population means while keeping the type I error low: ANOVA

2.- Between-groups design

- Assumptions:
  - The observations are random and independent samples from the populations
  - The distributions of the populations from which the samples are taken are normal
  - The variances of the distributions in the populations are equal (homoscedasticity)

Between-groups design

Example

### Table 14.2

<table>
<thead>
<tr>
<th>Method</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>6</td>
<td>10</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>6</td>
<td>11</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>6</td>
<td>11</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>6</td>
<td>11</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>6</td>
<td>11</td>
<td>11</td>
<td>6</td>
</tr>
</tbody>
</table>

- \( \alpha \) = level of significance of each separate test
- \( c \) = number of independent t tests

Between-groups design

Example

### Table 14.3

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>df</th>
<th>MV</th>
<th>MSE</th>
<th>( F )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between</td>
<td>110.83</td>
<td>4</td>
<td>27.71</td>
<td>11.20</td>
<td>2.16</td>
</tr>
<tr>
<td>Within</td>
<td>227.36</td>
<td>8</td>
<td>28.42</td>
<td>3.59</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>338.19</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Between-groups design

Example

### Table 14.4

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>df</th>
<th>MSE</th>
<th>( F )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between</td>
<td>110.83</td>
<td>4</td>
<td>27.71</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>227.36</td>
<td>8</td>
<td>28.42</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>338.19</td>
<td>12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** The MS for the total is not computed.

Therefore, the sample estimates of those values are

\[ F = \frac{MS_{between}}{MS_{within}} \]
3.- Effect size

- A most straightforward estimate:
  - Proportion of explained variance, $\eta^2$
  - $\eta^2 = \frac{SS_{\text{between}}}{SS_{\text{total}}}$
- Proportion of variance in the dependent variable that can be explained by or attributed to the independent variable

Cohen’s guidelines for $\eta^2$

<table>
<thead>
<tr>
<th>$\eta^2$</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>.01</td>
<td>Small</td>
</tr>
<tr>
<td>.09</td>
<td>Medium</td>
</tr>
<tr>
<td>.25</td>
<td>Large</td>
</tr>
</tbody>
</table>

In the case of the previous numerical example:
$\eta^2 = \frac{150.51}{237.94} = 0.63$

4.- Multiple comparisons

- The possible comparisons whenever more than two populations are involved
- As we saw already, $t$ test is not appropriate because it increases the probability of a type I error
- Tukey’s $HSD$ Test

Tukey’s $HSD$ Test

- $HSD$: ‘Honestly significant difference’ test
- Can be used to test all possible differences between pairs of means, and yet the cumulative probability of at least one type I error never exceeds the specified level of significance
  - $HSD = q \sqrt{\frac{MS_{\text{within}}}{n}}$
  - $n$: sample size in each group
  - $q$: ‘Studentized Range Statistic’ (Table G: $\alpha$, $k$, $df_{\text{within}}$)

- For the case of the previous numerical example with ($\alpha=.05$, $k=5$, $df_{\text{within}}=26$) $HSD = 4.17 \sqrt{\frac{3.36}{6.2}} = 3.0698$
- ($\alpha=.01$, $k=5$, $df_{\text{within}}=26$) $HSD = 5.17 \sqrt{\frac{3.36}{6.2}} = 3.806$

II. Inferential Statistics (10)

- Testing for a difference between means
- $t$ test for two related samples
1. Testing for a difference between means

- Two independent samples
  - Observations in one sample are not paired, on a one-to-one basis, with observations in the other sample

- Two related samples (repeated measures)
  - Observations in one sample are paired, on a one-to-one basis, with observations in the other sample

Testing for a difference between means (related samples)

- Statistical Hypotheses
  - \( H_0: \mu_D = 0 \)
  - \( H_1: \mu_D > 0 \)
  - \( H_0: \mu_D = 0 \)
  - \( H_1: \mu_D < 0 \)
  - \( H_0: \mu_D = 0 \)
  - \( H_1: \mu_D \neq 0 \)

Sampling distribution of \( \bar{D} \)

- Properties of the mean:
  \[ \mu_{\bar{D}} = \mu_D \]

2. \( t \) test for two related samples

- \( t \) ratio for two related samples
- ... as always:

\[ t = \frac{\bar{D} - \mu_{Dhyp}}{s_{\bar{D}}} \]

\( \sigma_D = \frac{\sigma_D}{\sqrt{n}} \)
1.- Repeated measures ANOVA

- A type of analysis that tests whether differences exist among population means with measures on the same subjects.

- Statistical Hypotheses
  - \( H_0: \mu_0 = \mu_1 = \mu_2 = \ldots = \mu_k \)
  - \( H_1: \) \( H_0 \) is false

Repeated measures ANOVA

- In One-way ANOVA:
  - \( SS_{total} = SS_{between} + SS_{within} \)

- In Repeated measures ANOVA:
  - \( SS_{within} = SS_{subject} + SS_{error} \)
  - \( SS_{total} = SS_{between} + SS_{subject} + SS_{error} \)

- This can yield a more powerful analysis

Sources of Variability

**Table 12.4**

<table>
<thead>
<tr>
<th>Formula</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( df_{total} )</td>
<td>Total degrees of freedom</td>
</tr>
<tr>
<td>( df_{between} )</td>
<td>Degrees of freedom between groups</td>
</tr>
<tr>
<td>( df_{within} )</td>
<td>Degrees of freedom within groups</td>
</tr>
</tbody>
</table>

**Figure 12.1**

Source of variability for one factor and repeated measures ANOVA
2.- Example

- See PDFs

HYPOTHESIS TEST SUMMARY
Repeated-Measures $F$ Test (Sleep Derivation Experiment)

Research Problem
On average, are subjects\' aggression scores in a controlled social situation affected by sleep deprivation periods of 0, 24, and 48 hours, where each subject experiences all three periods?

Statistical Hypotheses

$H_0: \mu_0 = \mu_{24} = \mu_{48}$

$H_1: \mu_i \neq \mu_j$ for some $i \neq j$

Decision Rule

Reject $H_0$ at the .05 level of significance if $F > 6.94$ (from Table C in Appendix C, given $df_{between} = 2$ and $df_{error} = 4$).

Calculations

$F = 27$ (See Tables 17.3 and 17.5 for more details.)

Decision

Reject $H_0$ at the .05 level of significance because $F = 27$ exceeds 6.94.

Interpretation

Mean aggression scores in a controlled social situation are affected by sleep deprivation when subjects experience all three levels of deprivation.

Table 17.5
ANOVA Table: Sleep Deprivation Experiment (Repeated Measures)

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between</td>
<td>54</td>
<td>2</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>Within</td>
<td>22</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subject</td>
<td>18</td>
<td>2</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>76</td>
<td>8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant at the .01 level.