Lecture 5: Comparing Treatment Means

Montgomery: Section 3-5
Linear Combination of Means

- ANOVA:

\[ y_{ij} = \mu + \tau_i + \epsilon_{ij} \]

\[ = \mu_i + \epsilon_{ij} \]

- Linear combination: \( L = c_1 \mu_1 + c_2 \mu_2 + \ldots + c_a \mu_a = \sum_{i=1}^{a} c_i \mu_i \)

- Want to test: \( H_0 : L = \sum c_i \mu_i = L_0 \)
  - pairwise comparison: \( \mu_i - \mu_j = 0 \)
  - treatment vs. control: \( \mu_i - \mu_1 = 0 \) if treatment 1 is control
  - comparing combinations of treatment effects: \( \mu_1 - 2\mu_2 + \mu_3 = 0 \).
  - high order effects (orthogonal polynomials): linear, quadratic, etc.
• Estimate of \( \hat{L} \):

\[
\hat{L} = \sum c_i \bar{y}_i.
\]

\[
\text{Var}(\hat{L}) = \sum c_i^2 \text{Var}(\bar{y}_i) = \sigma^2 \sum \frac{c_i^2}{n_i} \left( = \frac{\sigma^2}{n} \sum c_i^2 \right)
\]

• Standard Error for \( \hat{L} \)

\[
\text{S.E.}\hat{L} = \sqrt{\text{MSE} \sum \frac{c_i^2}{n_i}}
\]

• Test statistic

\[
t_0 = \frac{(\hat{L} - L_0)}{\text{S.E.}\hat{L}} \sim t(N - a) \text{ under } H_0
\]
Example: Lambs Diet Experiment

- Recall there are three diets and their treatment means are denoted by $\mu_1$, $\mu_2$ and $\mu_3$. Suppose one wants to consider

$$L = \mu_1 + 2\mu_2 + 3\mu_3 = 6\mu + \tau_1 + 2\tau_2 + 3\tau_3$$

and test $H_0 : L = 60$.

```plaintext
data lambs;
  input diet wtgain@@;
cards;
  1 8 1 16 1 9 2 9 2 16 2 21
  2 11 2 18 3 15 3 10 3 17 3 6;
proc glm;
class diet;
model wtgain=diet;
means diet;
estimate 'l1' intercept 6 diet 1 2 3;
run;
```
Example: Lambs Diet Experiment

- SAS output

<table>
<thead>
<tr>
<th>Level of diet</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>11.000</td>
<td>4.3589</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>15.000</td>
<td>4.9497</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>12.000</td>
<td>4.9666</td>
</tr>
</tbody>
</table>

Dependent Variable: wtgain

| Standard Parameter Estimate | Error   | t Value | Pr > |t| |
|----------------------------|---------|---------|------|---|
| l1                         | 77.000  | 8.885   | 8.67 | <.0001 |

- \(t_0 = (77.0 - 60) / 8.89 = 1.91\)

\[ P - \text{value} = P(t \leq -1.91 \text{ or } t \geq 1.91 | t(12 - 3)) = .088 \]

- Accept \(H_0: \mu_1 + 2\mu_2 + 3\mu_3 = 60\) at \(\alpha = 5\%\).
Contrasts

- \( \Gamma = \sum_{i=1}^{a} c_i \mu_i \) is a contrast if \( \sum_{i=1}^{a} c_i = 0 \).

  Equivalently, \( \Gamma = \sum_{i=1}^{a} c_i \tau_i \).

- Examples

  1. \( \Gamma_1 = \mu_1 - \mu_2 = \mu_1 - \mu_2 + 0\mu_3 + 0\mu_4 \),
     \( c_1 = 1, c_2 = -1, c_3 = 0, c_4 = 0 \)
     Comparing \( \mu_1 \) and \( \mu_2 \).

  2. \( \Gamma_2 = \mu_1 - 0.5\mu_2 - 0.5\mu_3 = \mu_1 - 0.5\mu_2 - 0.5\mu_3 + 0\mu_4 \)
     \( c_1 = 1, c_2 = -0.5, c_3 = -0.5, c_4 = 0 \)
     Comparing \( \mu_1 \) and the average of \( \mu_2 \) and \( \mu_3 \).

- Estimate for \( \Gamma \):

  \[
  C = \sum_{i=1}^{a} c_i \bar{y}_i.
  \]
Test: \( H_0 : \Gamma = 0 \)

\[
t_0 = \frac{C}{\text{S.E.}_C} \sim t(N - a)
\]

\[
t_0^2 = \frac{(\sum c_i \bar{y}_i.)^2}{\text{MSE} \sum \frac{c_i^2}{n_i}} = \frac{(\sum c_i \bar{y}_i.)^2 / \sum c_i^2 / n_i}{\text{MSE}} = \frac{SS_C / 1}{\text{MSE}}
\]

Under \( H_0 \), \( t_0^2 \sim F_{1, N-a} \).

• Contrast Sum of Squares

\[
SS_C = \frac{(\sum c_i \bar{y}_i.)^2}{\sum (c_i^2 / n_i)}
\]

\( SS_C \) is the amount of variation caused by \( \Gamma \).
SAS Code (cont.sas)
Tensile Strength Example

options ls=80;

title1 'Contrast Comparisons';

data one;
  infile 'c:saswork\data\tensile.dat';
  input percent strength time;

proc glm data=one;
  class percent;
  model strength=percent;
  contrast 'C1' percent 0 0 0 1 -1;
  contrast 'C2' percent 1 0 1 -1 -1;
  contrast 'C3' percent 1 0 -1 0 0;
  contrast 'C4' percent 1 -4 1 1 1;
### Dependent Variable: STRENGTH

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Squares</th>
<th>Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>4</td>
<td>475.76000</td>
<td>118.94000</td>
<td>14.76</td>
<td>0.0001</td>
</tr>
<tr>
<td>Error</td>
<td>20</td>
<td>161.20000</td>
<td>8.06000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>24</td>
<td>636.96000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type I SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>PERCENT</td>
<td>4</td>
<td>475.76000</td>
<td>118.94000</td>
<td>14.76</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Contrast</th>
<th>DF</th>
<th>Contrast SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>1</td>
<td>291.60000</td>
<td>291.60000</td>
<td>36.18</td>
<td>0.0001</td>
</tr>
<tr>
<td>C2</td>
<td>1</td>
<td>31.25000</td>
<td>31.25000</td>
<td>3.88</td>
<td>0.0630</td>
</tr>
<tr>
<td>C3</td>
<td>1</td>
<td>152.10000</td>
<td>152.10000</td>
<td>18.87</td>
<td>0.0003</td>
</tr>
<tr>
<td>C4</td>
<td>1</td>
<td>0.81000</td>
<td>0.81000</td>
<td>0.10</td>
<td>0.7545</td>
</tr>
</tbody>
</table>
Orthogonal Contrasts

- Two contrasts \( \{c_i\} \) and \( \{d_i\} \) are **Orthogonal** if
  \[
  \sum_{i=1}^{a} \frac{c_id_i}{n_i} = 0
  \]

- Example
  \[
  \Gamma_1 = \mu_1 + \mu_2 - \mu_3 - \mu_4, \quad \text{So} \quad c_1 = 1, c_2 = 1, c_3 = -1, c_4 = -1.
  \]
  \[
  \Gamma_2 = \mu_1 - \mu_2 + \mu_3 - \mu_4. \quad \text{So} \quad d_1 = 1, d_2 = -1, d_3 = 1, d_4 = -1
  \]
  It is easy to verify that both \( \Gamma_1 \) and \( \Gamma_2 \) are contrasts. Furthermore,
  \[
  c_1d_1 + c_2d_2 + c_3d_3 + c_4d_4 =
  1 \times 1 + 1 \times (-1) + (-1) \times 1 + (-1) \times (-1) = 0.
  \]
  Hence, \( \Gamma_1 \) and \( \Gamma_2 \) are orthogonal to each other.

- A **complete set** of orthogonal contrasts \( C = \{\Gamma_1, \Gamma_2, \ldots, \Gamma_{a-1}\} \) if contrasts are mutually orthogonal and there does not exist a contrast orthogonal to all contrasts in \( C \).
• If there are \( a \) treatments, \( C \) must contain \( a - 1 \) contrasts.

• Complete set is not unique. For example, in the tensile strength example

\[
\Gamma_1 = (0, 0, 0, 1, -1) \\
\Gamma_2 = (1, 0, 1, -1, -1) \\
\Gamma_3 = (1, 0, -1, 0, 0) \\
\Gamma_4 = (1, -4, 1, 1, 1)
\]

\( C_1 \) : includes:

\[
\Gamma'_1 = (-2, -1, 0, 1, 2) \\
\Gamma'_2 = (2, -1, -2, -1, 2) \\
\Gamma'_3 = (-1, 2, 0, -2, 1) \\
\Gamma'_4 = (1, -4, 6, -4, 1)
\]

\( C_2 \) : includes:
Orthogonal Contrasts

- Orthogonal contrasts (estimates) are independent with each other.
- Suppose $C_1, C_2, \ldots, C_{a-1}$ are the estimates of the contrasts in a complete set of contrasts $\{\Gamma_1, \Gamma_2, \ldots, \Gamma_{a-1}\}$, then

$$SS_{\text{Treatment}} = SS_{C_1} + SS_{C_2} + \cdots + SS_{C_{a-1}}$$

- Recall in ANOVA, $F_0 = \frac{\text{MS}_{\text{Treatment}}}{\text{MSE}}$,

$$F_0 = \frac{SS_{C_1}/\text{MSE} + \cdots + SS_{C_{a-1}}/\text{MSE}}{a-1} = \frac{F_{10} + F_{20} + \cdots + F_{(a-1)0}}{a-1}$$

where $F_{i0}$ is the test statistic used to test contrast $\Gamma_i$.

- Example on Slide 9
Tensile Example

Try to model mean response as a function of treatments
Orthogonal contrasts and orthogonal polynomial model

- Treatments are quantitative (assume \( a = 4 \))
- One can use general polynomial model to fit the trend (\( t \): level or treatment).

\[
f(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3
\]

Regression can be used to get the estimates for \( a_1 \), \( a_2 \) and \( a_3 \).

- We will use orthogonal polynomial model

\[
f(t) = \beta_0 + \beta_1 P_1(t) + \beta_2 P_2(t) + \beta_3 P_3(t)
\]

where \( P_1(t) \), \( P_2(t) \) and \( P_3(t) \) are pre-specified polynomials of order 1, 2 and 3, respectively. \( P_1(t) \) is linear, \( P_2(t) \) is quadratic and \( P_3(t) \) is cubic.

Let \( t_1, t_2, \ldots, t_a \) are the treatments (equally spaced), then the polynomials corresponds to the following contrasts:
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Contrasts</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$</td>
<td>$t_1$</td>
<td>$t_2$</td>
<td>$\cdots$</td>
<td>$t_a$</td>
<td>$\Gamma_1$</td>
<td>$D_1$</td>
</tr>
<tr>
<td>$P_1(t)$</td>
<td>$P_1(t_1)$</td>
<td>$P_1(t_2)$</td>
<td>$\cdots$</td>
<td>$P_1(t_a)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$P_2(t)$</td>
<td>$P_2(t_1)$</td>
<td>$P_2(t_2)$</td>
<td>$\cdots$</td>
<td>$P_2(t_a)$</td>
<td>$\Gamma_2$</td>
<td>$D_2$</td>
</tr>
<tr>
<td>$P_3(t)$</td>
<td>$P_3(t_1)$</td>
<td>$P_3(t_2)$</td>
<td>$\cdots$</td>
<td>$P_3(t_a)$</td>
<td>$\Gamma_3$</td>
<td>$D_3$</td>
</tr>
</tbody>
</table>

where

$$D_i = P_i(t_1)^2 + P_i(t_2)^2 + \cdots + P_i(t_a)^2$$

If $\Gamma_1$, $\Gamma_2$ and $\Gamma_3$ are orthogonal to each other, then we say $P_1(t)$, $P_2(t)$ and $P_3(t)$ are orthogonal polynomials.

- Coefficients $\beta_i$ can be estimated and tested by the contrasts $\Gamma_1$, $\Gamma_2$ and $\Gamma_3$.
- Predict $f(t)$ when $t$ is not a treatment used in the experiment.
tensile strength example: orthogonal polynomial effects

- Treatment levels $t_k$: 15, 20, 25, 30, 35; Median: 25; Pace: 5
- Orthogonal polynomials: let $x = (t - 25)/5$.

$$
P_1(t) = x
$$
$$
P_2(t) = x^2 - 2
$$
$$
P_3(t) = \frac{5}{6}[x^3 - 17x/5]
$$
$$
P_4(t) = \frac{35}{12}[x^4 - 31x/7 + 72/35]
$$

- Polynomial Contrasts and Effects

<table>
<thead>
<tr>
<th>$t$</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
<th>35</th>
<th>Contrast</th>
<th>$D$</th>
<th>Effect (Trend)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_1(t)$</td>
<td>-2</td>
<td>-1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>$\Gamma_1$</td>
<td>$D_1 = 10$</td>
<td>linear</td>
</tr>
<tr>
<td>$P_2(t)$</td>
<td>2</td>
<td>-1</td>
<td>-2</td>
<td>-1</td>
<td>2</td>
<td>$\Gamma_2$</td>
<td>$D_2 = 14$</td>
<td>quadratic</td>
</tr>
<tr>
<td>$P_3(t)$</td>
<td>-1</td>
<td>2</td>
<td>0</td>
<td>-2</td>
<td>1</td>
<td>$\Gamma_3$</td>
<td>$D_3 = 10$</td>
<td>cubic</td>
</tr>
<tr>
<td>$P_4(t)$</td>
<td>1</td>
<td>-4</td>
<td>6</td>
<td>-4</td>
<td>1</td>
<td>$\Gamma_4$</td>
<td>$D_4 = 70$</td>
<td>4th order</td>
</tr>
</tbody>
</table>

- The contrasts can be directly derived from Table IX or Table X.
Want to fit the model

\[ f(t) = \beta_0 + \beta_1 P_1(t) + \beta_2 P_2(t) + \beta_3 P_3(t) + \beta_4 P_4(t) \]

Estimation and Testing
- \( \beta_1 \): use \( \Gamma_1 \),

\[ \hat{\beta}_1 = \frac{c_{11} \bar{y}_1. + \cdots + c_{15} \bar{y}_5.}{D_1} \]

Test: \( H_0 : \beta_1 = 0, F_{10} = \frac{SS_{C_i}}{MSE} \sim F_{1,N-5} \).
- \( \beta_2 \): use \( \Gamma_2 \),

\[ \hat{\beta}_2 = \frac{c_{21} \bar{y}_1. + \cdots + c_{25} \bar{y}_5.}{D_2} \]

Test: \( H_0 : \beta_2 = 0, F_{20} = \frac{SS_{C_i}}{MSE} \sim F_{1,N-5} \)
- Similar for \( \beta_3 \) and \( \beta_4 \)

Question: what is the estimate for \( \beta_0 \)?
General formulas for orthogonal polynomial of degrees 1-4

One factor of $a$ levels $l_1, l_2, \ldots, l_a$, equally spaced. Let $m$ be the median, $\delta$ be the difference between two consecutive levels:

$$P_1(t) = \lambda_1 \left( \frac{t-m}{\delta} \right)$$

$$P_2(t) = \lambda_2 \left[ \left( \frac{t-m}{\delta} \right)^2 - \frac{a^2 - 1}{12} \right]$$

$$P_3(t) = \lambda_3 \left[ \left( \frac{t-m}{\delta} \right)^3 - \left( \frac{t-m}{\delta} \right) \left( \frac{3a^2 - 7}{20} \right) \right]$$

$$P_4(t) = \lambda_4 \left[ \left( \frac{t-m}{\delta} \right)^4 - \left( \frac{t-m}{\delta} \right)^2 \left( \frac{3a^2 - 13}{14} \right) + \frac{3(a^2 - 1)(a^2 - 9)}{560} \right]$$

$(\lambda_i)$ are constants to make the polynomials have integer values at the treatment levels, they are available from Table IX or Table X.

Tensile Strength Example: $m=25$, $\delta = 5$, $(\lambda_i) = (1, 1, 5/6, 35/12)$
SAS
tensile strength example

data one;
  infile 'c:saswork\data\tensile.dat';
  input percent strength time;

proc glm data=one;
  class percent;
  model strength=percent;
  estimate 'C1' percent -2 -1  0  1  2;
  estimate 'C2' percent  2 -1 -2 -1  2;
  estimate 'C3' percent -1  2  0 -2  1;
  estimate 'C4' percent  1 -4  6 -4  1;
  contrast 'C1' percent -2 -1  0  1  2;
  contrast 'C2' percent  2 -1 -2 -1  2;
  contrast 'C3' percent -1  2  0 -2  1;
  contrast 'C4' percent  1 -4  6 -4  1;
run;
## Output

**Dependent Variable:** strength

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>4</td>
<td>475.7600000</td>
<td>118.9400000</td>
<td>14.76</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>20</td>
<td>161.2000000</td>
<td>8.0600000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>24</td>
<td>636.9600000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Parameter | Estimate | Error | t Value | Pr > |t| |
|-----------|----------|-------|---------|------|---|
| C1        | 8.2000000| 4.0149720| 2.04   | 0.0545|
| C2        | -31.0000000| 4.7505789| -6.53  | <.0001|
| C3        | -11.4000000| 4.0149720| -2.84  | 0.0101|
| C4        | -21.8000000| 10.6226174| -2.05  | 0.0535|

<table>
<thead>
<tr>
<th>Contrast</th>
<th>DF</th>
<th>Contrast SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>1</td>
<td>33.62000000</td>
<td>33.62000000</td>
<td>4.17</td>
<td>0.0545</td>
</tr>
<tr>
<td>C2</td>
<td>1</td>
<td>343.2142857</td>
<td>343.2142857</td>
<td>42.58</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>C3</td>
<td>1</td>
<td>64.98000000</td>
<td>64.98000000</td>
<td>8.06</td>
<td>0.0101</td>
</tr>
<tr>
<td>C4</td>
<td>1</td>
<td>33.9457143</td>
<td>33.9457143</td>
<td>4.21</td>
<td>0.0535</td>
</tr>
</tbody>
</table>
Estimates

Hence,

\[ \hat{\beta}_1 = \frac{8.20}{10} = 0.82; \quad \hat{\beta}_2 = \frac{-31}{14} = -2.214 \]

\[ \hat{\beta}_3 = \frac{-11.4}{10} = -1.14; \quad \hat{\beta}_4 = \frac{-21.8}{70} = -0.311 \]

So the fitted functional relationship between tensile strength \( y \) and cotton percent \( (t) \) is

\[ y = \hat{\beta}_0 + 0.82P_1(t) - 2.214P_2(t) - 1.14P_3(t) - 0.311P_4(t), \]

where \( P_1(t), \ldots, P_4(t) \) are defined on Slide 16.
Testing Multiple Contrasts (Multiple Comparisons) Using Confidence Intervals

- One contrast:

\[ H_0 : \Gamma = \sum c_i \mu_i = \Gamma_0 \text{ vs } H_1 : \Gamma \neq \Gamma_0 \text{ at } \alpha \]

100(1-\alpha) Confidence Interval (CI) for \( \Gamma \):

\[ \text{CI} : \sum c_i y_i. \pm t_{\alpha/2, N-a} \sqrt{MSE \sum \frac{c_i^2}{n_i}} \]

\[ P(\text{CI not contain } L_0 \mid H_0) = \alpha (= \text{type I error}) \]

- Decision Rule: Reject \( H_0 \) if CI does not contain \( \Gamma_0 \).
• Multiple contrasts

\[ H_0 : \Gamma^1 = \Gamma^1_0, \ldots \Gamma^m = \Gamma^m_0 \text{ vs } H_1 : \text{at least one does not hold} \]

If we construct CI_1, CI_2, ..., CI_m, each with 100(1-\(\alpha\)) level, then for each CI_i,

\[ P(\text{CI}_i \text{ not contain } \Gamma^i_0 \mid H_0) = \alpha, \text{ for } i = 1, \ldots, m \]

• But the overall error rate (probability of type I error for \(H_0\) vs \(H_1\)) is inflated and much larger than \(\alpha\), that is,

\[ P(\text{at least one CI}_i \text{ not contain } \Gamma^i_0 \mid H_0) >> \alpha \]

• One way to achieve small overall error rate, we require much smaller error rate (\(\alpha'\)) of each individual CI_i.
Bonferroni Method for Testing Multiple Contrasts

- Bonferroni Inequality

\[ P( \text{at least one } CI_i \text{ not contain } \Gamma_0^i \mid H_0) \]

\[ = P(\text{CI}_1 \text{ not contain..or } \ldots \text{or } \text{CI}_m \text{ not contain } \mid H_0) \]

\[ \leq P(\text{CI}_1 \text{ not } \mid H_0) + \ldots + P(\text{CI}_m \text{ not } \mid H_0) = m\alpha' \]

- In order to control overall error rate (or, overall confidence level), let

\[ m\alpha' = \alpha, \text{ we have, } \alpha' = \alpha/m \]

- Bonferroni CIs:

\[ CI_i : \sum c_{ij} \bar{y}_j. \pm t_{\alpha/2m}(N-a) \sqrt{MS_E} \sum \frac{c^2_{ij}}{n_j} \]

- When \( m \) is large, Bonferroni CIs are too conservative (overall type II error too large).
Scheffe’s method for Testing All Contrasts

- Consider all possible contrasts: \( \Gamma = \sum c_i \mu_i \)
  
  Estimate: \( C = \sum c_i \bar{y}_i \),  
  St. Error: \( \text{S.E.}_C = \sqrt{\text{MSE} \sum \frac{c_i^2}{n_i}} \)

- Critical value: \( \sqrt{(a - 1)F_{\alpha, a-1, N-a}} \)

- Scheffe’s simultaneous CI: \( C \pm \sqrt{(a - 1)F_{\alpha, a-1, N-a}} \text{ S.E.}_C \)

- Overall confidence level and error rate for \( m \) contrasts

\[ P(\text{CIs contain true parameter for any contrast}) \geq 1 - \alpha \]

\[ P(\text{at least one CI does not contain true parameter}) \leq \alpha \]

Remark: Scheffe’s method is also conservative, too conservative when \( m \) is small
Methods for Pairwise Comparisons

- There are \( a(a - 1)/2 \) possible pairs: \( \mu_i - \mu_j \) (contrast for comparing \( \mu_i \) and \( \mu_j \)). We may be interested in \( m \) pairs or all pairs.

- Standard Procedure:
  1. Estimation: \( \bar{y}_i - \bar{y}_j \).
  2. Compute a **Critical Difference** (CD) (based on the method employed)
  3. If

\[
| \bar{y}_i - \bar{y}_j | > CD
\]

or equivalently if the interval

\[
(\bar{y}_i - \bar{y}_j - CD, \ \bar{y}_i - \bar{y}_j + CD)
\]

does not contain zero, declare \( \mu_i - \mu_j \) significant.
Methods for Calculating CD.

- Least significant difference (LSD):
  \[
  CD = t_{\alpha/2, N-a} \sqrt{MS_E (1/n_i + 1/n_j)}
  \]
  not control overall error rate

- Bonferroni method (for \(m\) pairs)
  \[
  CD = t_{\alpha/2m, N-a} \sqrt{MS_E (1/n_i + 1/n_j)}
  \]
  control overall error rate for the \(m\) comparisons.

- Tukey’s method (for all possible pairs)
  \[
  CD = \frac{q_{\alpha} (a, N-a)}{\sqrt{2}} \sqrt{MS_E (1/n_i + 1/n_j)}
  \]
  \(q_{\alpha} (a, N-a)\) from studentized range distribution (Table VII or Table VIII).
  Control overall error rate (exact for balanced experiments). (Example 3.7).
Comparing treatments with control (Dunnett’s method)

1. Assume $\mu_1$ is a control, and $\mu_2, \ldots, \mu_a$ are (new) treatments

2. Only interested in $a - 1$ pairs: $\mu_2 - \mu_1, \ldots, \mu_a - \mu_1$

3. Compare $|\bar{y}_i - \bar{y}_1|$ to

$$CD = d_\alpha(a - 1, N - a) \sqrt{MS_E(1/n_i + 1/n_1)}$$

where $d_\alpha(p, f)$ from Table IX or Table VIII: critical values for Dunnett’s test.

4. Remark: control overall error rate. Read Example 3-9 (or 3-10)

For pairwise comparison, which method should be preferred? LSD, Bonferroni, Tukey, Dunnett or others?
SAS Code

data one;
  infile 'c:saswork\data\tensile.dat';
  input percent strength time;

proc glm data=one;
  class percent;
  model strength=percent;

  /* Construct CI for Treatment Means*/
  means percent /alpha=.05 lsd clm;
  means percent / alpha=.05 bon clm;

  /* Pairwise Comparison*/
  means percent /alpha=.05 lines lsd;
  means percent /alpha=.05 lines bon;
  means percent /alpha=.05 lines scheffe;
  means percent /alpha=.05 lines tukey;
  means percent /dunnett;
run;
The GLM Procedure

t Confidence Intervals for \( y \)

<table>
<thead>
<tr>
<th>Alpha</th>
<th>0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error Degrees of Freedom</td>
<td>20</td>
</tr>
<tr>
<td>Error Mean Square</td>
<td>8.06</td>
</tr>
<tr>
<td>Critical Value of ( t )</td>
<td>2.08596</td>
</tr>
<tr>
<td>Half Width of Confidence Interval</td>
<td>2.648434</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>trt</th>
<th>N</th>
<th>Mean</th>
<th>Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>5</td>
<td>21.600</td>
<td>18.952 24.248</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>17.600</td>
<td>14.952 20.248</td>
</tr>
<tr>
<td>20</td>
<td>5</td>
<td>15.400</td>
<td>12.752 18.048</td>
</tr>
<tr>
<td>35</td>
<td>5</td>
<td>10.800</td>
<td>8.152 13.448</td>
</tr>
<tr>
<td>15</td>
<td>5</td>
<td>9.800</td>
<td>7.152 12.448</td>
</tr>
</tbody>
</table>
The GLM Procedure

Bonferroni t Confidence Intervals for $y$

Alpha 0.05
Error Degrees of Freedom 20
Error Mean Square 8.06
Critical Value of t 2.84534
Half Width of Confidence Interval 3.612573

<table>
<thead>
<tr>
<th>trt</th>
<th>N</th>
<th>Mean</th>
<th>Simultaneous 95% Confidence Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>5</td>
<td>21.600</td>
<td>17.987 25.213</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>17.600</td>
<td>13.987 21.213</td>
</tr>
<tr>
<td>20</td>
<td>5</td>
<td>15.400</td>
<td>11.787 19.013</td>
</tr>
<tr>
<td>35</td>
<td>5</td>
<td>10.800</td>
<td>7.187 14.413</td>
</tr>
<tr>
<td>15</td>
<td>5</td>
<td>9.800</td>
<td>6.187 13.413</td>
</tr>
</tbody>
</table>
t Tests (LSD) for y
NOTE: This test controls the Type I comparisonwise error rate, not the experimentwise error rate.

Alpha 0.05
Error Degrees of Freedom 20
Error Mean Square 8.06
Critical Value of t 2.08596
Least Significant Difference 3.7455

Means with the same letter are not significantly different.

<table>
<thead>
<tr>
<th>t Grouping</th>
<th>Mean</th>
<th>N</th>
<th>trt</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>21.600</td>
<td>5</td>
<td>30</td>
</tr>
<tr>
<td>B</td>
<td>17.600</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>B</td>
<td>15.400</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>C</td>
<td>10.800</td>
<td>5</td>
<td>35</td>
</tr>
<tr>
<td>C</td>
<td>9.800</td>
<td>5</td>
<td>15</td>
</tr>
</tbody>
</table>
Bonferroni (Dunn) t Tests for y

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha 0.05
Error Degrees of Freedom 20
Error Mean Square 8.06
Critical Value of t 3.15340
Minimum Significant Difference 5.6621

Means with the same letter are not significantly different.

<table>
<thead>
<tr>
<th>Bon Grouping</th>
<th>Mean</th>
<th>N</th>
<th>trt</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>21.600</td>
<td>5</td>
<td>30</td>
</tr>
<tr>
<td>B</td>
<td>17.600</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>C</td>
<td>15.400</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>10.800</td>
<td>5</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>9.800</td>
<td>5</td>
<td>15</td>
</tr>
</tbody>
</table>
Scheffe’s Test for $\gamma$

NOTE: This test controls the Type I experimentwise error rate.

Alpha 0.05
Error Degrees of Freedom 20
Error Mean Square 8.06
Critical Value of $F$ 2.86608
Minimum Significant Difference 6.0796

Means with the same letter are not significantly different.

<table>
<thead>
<tr>
<th>Scheffe Grouping</th>
<th>Mean</th>
<th>$N$</th>
<th>trt</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>21.600</td>
<td>5</td>
<td>30</td>
</tr>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>A</td>
<td>17.600</td>
<td>5</td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>C</td>
<td>15.400</td>
<td>5</td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>9.800</td>
<td>5</td>
<td>15</td>
</tr>
</tbody>
</table>
Tukey’s Studentized Range (HSD) Test for $y$

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

<table>
<thead>
<tr>
<th>Alpha</th>
<th>0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error Degrees of Freedom</td>
<td>20</td>
</tr>
<tr>
<td>Error Mean Square</td>
<td>8.06</td>
</tr>
<tr>
<td>Critical Value of Studentized Range</td>
<td>4.23186</td>
</tr>
<tr>
<td>Minimum Significant Difference</td>
<td>5.373</td>
</tr>
</tbody>
</table>

Means with the same letter are not significantly different.

<table>
<thead>
<tr>
<th>Tukey Grouping</th>
<th>Mean</th>
<th>N</th>
<th>trt</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>21.600</td>
<td>5</td>
<td>30</td>
</tr>
<tr>
<td>B</td>
<td>17.600</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>B</td>
<td>15.400</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>C</td>
<td>10.800</td>
<td>5</td>
<td>35</td>
</tr>
<tr>
<td>D</td>
<td>9.800</td>
<td>5</td>
<td>15</td>
</tr>
</tbody>
</table>
Dunnett’s t Tests for $y$

NOTE: This test controls the Type I experimentwise error for comparisons of treatments against a control.

<table>
<thead>
<tr>
<th>Alpha</th>
<th>0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error Degrees of Freedom</td>
<td>20</td>
</tr>
<tr>
<td>Error Mean Square</td>
<td>8.06</td>
</tr>
<tr>
<td>Critical Value of Dunnett’s t</td>
<td>2.65112</td>
</tr>
<tr>
<td>Minimum Significant Difference</td>
<td>4.7602</td>
</tr>
</tbody>
</table>

Comparisons significant at the 0.05 level are indicated by ***.

<table>
<thead>
<tr>
<th>Difference</th>
<th>Between Means</th>
<th>Simultaneous 95% Confidence Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>trt Comparison</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30 - 15</td>
<td>11.800</td>
<td>7.040 16.560 ***</td>
</tr>
<tr>
<td>25 - 15</td>
<td>7.800</td>
<td>3.040 12.560 ***</td>
</tr>
<tr>
<td>20 - 15</td>
<td>5.600</td>
<td>0.840 10.360 ***</td>
</tr>
<tr>
<td>35 - 15</td>
<td>1.000</td>
<td>-3.760 5.760</td>
</tr>
</tbody>
</table>
Determining Sample Size

- More replicates required to detect small treatment effects
- Operating Characteristic Curves for $F$ tests
- Probability of type II error

$$
\beta = P( \text{accept } H_0 \mid H_0 \text{ is false})
$$

$$
= P(F_0 < F_{\alpha, a-1, N-a} \mid H_1 \text{ is correct })
$$

- Under $H_1$, $F_0$ follows a noncentral $F$ distribution with noncentrality $\lambda$ and degrees of freedom, $a - 1$ and $N - a$. Let

$$
\Phi^2 = \frac{n \sum_{i=1}^{a} \tau_i^2}{a \sigma^2}
$$

- OC curves of $\beta$ vs $n$ and $\Phi$ are included in Chart V for various $\alpha$ and $a$.
- Read Example 3.10 or 3.11.