Section 9.2

- The overall F-test for the fishing line example just tells us that not all means are equal.
- We usually want more detailed information comparing particular groups.

Fishing line example  Stren    Triline XL    Triline XT

\[ \mu_1 \quad \mu_2 \quad \mu_3 \]

We might be interested in

- \( \mu_1 - \mu_2 \)  Stren vs XL
- \( \mu_1 - \mu_3 \)  Stren vs XT
- \( \mu_2 - \mu_3 \)  XL vs XT
- \( \mu_1 - \frac{\mu_2 + \mu_3}{2} \)  Stren vs Trilene

Each of these comparisons is a linear combination of means, denoted \( l \).

\[
\begin{align*}
l = \sum_{i=1}^{t} a_i \mu_i &= a_1 \mu_1 + a_2 \mu_2 + a_3 \mu_3 \\
\end{align*}
\]

For example

\[
\begin{align*}
\mu_1 - \frac{\mu_2 + \mu_3}{2} &= 1 \mu_1 - \frac{1}{2} \mu_2 - \frac{1}{2} \mu_3 \\
\end{align*}
\]

\[
\begin{align*}
a_1 &= 1 \quad a_2 = -\frac{1}{2} \quad a_3 = -\frac{1}{2}
\end{align*}
\]

A contrast has \( \sum a_i = 0 \)

The estimate for \( l \) is \( \hat{l} = \sum a_i \bar{y}_i = a_1 \bar{y}_1 + a_2 \bar{y}_2 + a_3 \bar{y}_3 \).

\[
\begin{align*}
\text{Var}(\hat{l}) &= V(\hat{l}) \\
&= a_1^2 \text{Var}(\bar{y}_1) + a_2^2 \text{Var}(\bar{y}_2) + a_3^2 \text{Var}(\bar{y}_3) \\
&= a_1^2 \frac{\sigma^2}{n_1} + a_2^2 \frac{\sigma^2}{n_2} + a_3^2 \frac{\sigma^2}{n_3} = \sigma^2 \sum_{i=1}^{t} \frac{a_i^2}{n_i}
\end{align*}
\]

The estimated variance is \( \hat{\text{Var}}(\hat{l}) = s_w^2 \sum_{i=1}^{t} \frac{a_i^2}{n_i} = MS_{\text{Error}} \sum_{i=1}^{t} \frac{a_i^2}{n_i} \)

\[ df = df_{\text{Error}} \]
A confidence interval for $l$ is

$$\hat{l} \pm t_{\alpha/2} \sqrt{V(\hat{l})}$$

$$\hat{l} \pm t_{\alpha/2} SE_l$$

To test $H_0 : l = 0$  

$$t = \frac{\hat{l} - 0}{SE_l}$$

Sometimes contrasts are tested with an equivalent F-test.

$$t_{df}^2 = F_{1,df}^2$$

$$F = t^2 = \frac{\bar{l}^2}{\hat{V}(\bar{l})} = \frac{\bar{l}^2}{s^2_w \sum(a_i^2 / n_i)} = \frac{\bar{l}^2 / \sum(a_i^2 / n_i)}{s^2_w} = \frac{MS_{Contrast}}{MS_{Error}}$$

The df for a single contrast is 1.

The SS for a contrast is $SS_{Contrast} = MS_{Contrast} = \frac{\bar{l}^2}{\sum(a_i^2 / n_i)}$

Two contrasts:

$$l_1 = \sum a_i \mu_i \quad l_2 = \sum b_i \mu_i$$ are orthogonal (⊥) if $\sum \frac{a_i b_i}{n_i} = 0$

For equal $n_i$, the contrasts are orthogonal if $\sum a_i b_i = 0$ just like perpendicular vectors.

<table>
<thead>
<tr>
<th></th>
<th>XL vs XT</th>
<th>Stren vs Tri</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stren</td>
<td>0</td>
<td>0 × 1 = 0</td>
</tr>
<tr>
<td>Tri-XL</td>
<td>1</td>
<td>1 × (−1/2) = −1/2</td>
</tr>
<tr>
<td>Tri-XT</td>
<td>−1</td>
<td>(−1) × (−1/2) = +1/2</td>
</tr>
</tbody>
</table>

- Sum = 0

- Since the sum is 0, these contrasts are orthogonal.
- The 2 df for treatment effects can be viewed as two 1 df orthogonal contrasts
  - $SS_{treat} = SS_{Contrast_1} + SS_{Contrast_2}$
- Testing $\mu_1 = \mu_2 = \mu_3$ is equivalent to testing that both of the following comparisons are true or that these contrasts equal zero.
  - $\mu_2 = \mu_3$  \quad $\mu_1 = \frac{\mu_2 + \mu_3}{2}$
  - $\mu_2 - \mu_3 = 0$  \quad $\mu_1 - \frac{\mu_2 + \mu_3}{2} = 0$
- Equality of $t$ treatments is equivalent to $t − 1$ linearly independent contrast all being equal to zero.
- This concept of breaking treatment sums of squares into orthogonal components will be important later in more complicated situations in chapter 15.
To find $SS_{C_1}$ and $SS_{C_2}$, it’s easier to work with vectors without fractions.

\[ \mu_2 - \mu_3 = 0 \quad 2\mu_1 - \mu_2 - \mu_3 = 0 \]

\[
\begin{array}{cc}
C_1 & C_2 \\
\hline
\text{Stren} & 0 & 2 \\
\text{XL} & 1 & -1 \\
\text{XT} & -1 & 1 \\
\end{array}
\]

\[
SS_{C_1} = \frac{\hat{\ell}^2}{\sum(a_i^2/n_i)} = \frac{(11.52 - 11.64)^2}{\frac{1^2}{10} + \frac{1^2}{10}} = 0.072
\]

\[
SS_{C_2} = \frac{(2(11.1) - 1(11.5) - 1(11.64))^2}{\frac{2^2}{10} + \frac{(-1)^2}{10} + \frac{(-1)^2}{10}} = 1.536
\]

\[
SS_{C_1} \quad 0.072 \\
SS_{C_2} \quad 1.536 \\
SSB \quad 1.608
\]

To test $H_0 : \mu_2 - \mu_3 = 0$, we have options

- t-test & confidence interval
- F-test

Confidence Interval

\[ \hat{\ell} = \bar{y}_2 - \bar{y}_3 = 11.52 - 11.64 = -0.12 \]

\[ \hat{\ell} = MS_{\text{Error}} \sum a_i^2/n = 0.01556 \left( \frac{1^2}{10} + \frac{(-1)^2}{10} \right) = 0.00311 \]

\[ SE_{\hat{\ell}} = \sqrt{\hat{\ell}} = \sqrt{0.00311} = 0.0558 \]

\[ -0.12 \pm 2.052(0.0558) \]

\[ -0.12 \pm 0.114 \]

\[ -0.234 \text{ to } -0.066 \]

Since 0 is not in the confidence interval,

- we reject $H_0 : \mu_2 - \mu_3 = 0$ vs $H_a : \mu_2 - \mu_3 \neq 0$
- at the $\alpha = 0.05$ level.
Testing with the t-test

\[ t = \frac{-0.12}{0.0558} = -2.15 \]

\[ 2 \times 0.01 < p\text{-value} < 2 \times 0.025 \]

\[ 0.02 < p\text{-value} < 0.05 \]

Equivalently

\[ F = \frac{SS_{C1} / 1}{MS_{\text{Error}}} \]

\[ F = \frac{0.072}{0.0156} = 4.63 = (-2.15)^2 \]

\[ 0.025 < p < 0.05 \]

Another example \( C_2 = \mu_1 - \frac{\mu_2 + \mu_3}{2} \)

\[ \hat{l} = 11.10 - \frac{11.52 + 11.64}{2} = -0.48 \]

\[ SE_i = \sqrt{MSE \sum \frac{a_i^2}{n_i}} = \sqrt{0.01556 \left( \frac{12}{10} + \frac{(-1/2)^2}{10} + \frac{(-1/2)^2}{10} \right)} \]

\[ = \sqrt{0.01556 \left( \frac{4}{40} + \frac{1}{40} + \frac{1}{40} \right)} = \sqrt{0.00233} = 0.0483 \]

\[ -0.48 \pm 2.052(0.0483) \]

\[ -0.48 \pm 0.099 \]

\[ -0.579 \text{ to } -0.381 \]

\[ t = \frac{-0.48}{0.0483} = -9.94 \]

\[ p\text{-value} < 2 \times 0.0005 \]

\[ p\text{-value} < 0.001 \]

\[ F = \frac{SS_{C2} / 1}{MS_{\text{Error}}} = \frac{1.536}{0.0156} = 98.74 \]

\[ p < 0.001 \]

*Note: \( F = 98.74 = (9.94)^2 \)*
Aside (means not on any test): The reason we need $\sum \frac{a_i b_i}{n_i} = 0$ for orthogonal contrasts is so that the vectors of coefficients of the individual $y_{ij}$ values are orthogonal.

$$\sum a_i \bar{y}_i. = \sum a_i \frac{\sum y_{ij}}{n_i} = \sum \frac{a_i}{n_i} y_{ij}$$

$$\begin{align*}
y_{11} & \quad a_1/n_1 & \quad b_1/n_1 & \quad a_1 b_1/n_1^2 \\
y_{12} & \quad a_1/n_1 & \quad b_1/n_1 & \quad a_1 b_1/n_1^2 \\
\vdots & \quad \vdots & \quad \vdots & \quad \vdots \\
y_{1n_1} & \quad a_1/n_1 & \quad b_1/n_1 & \quad a_1 b_1/n_1^2 \\
\vdots & \quad \vdots & \quad \vdots & \quad \vdots \\
y_{t1} & \quad a_t/n_t & \quad b_t/n_t & \quad a_t b_t/n_t^2 \\
y_{t2} & \quad a_t/n_t & \quad b_t/n_t & \quad a_t b_t/n_t^2 \\
\vdots & \quad \vdots & \quad \vdots & \quad \vdots \\
y_{tn_t} & \quad a_t/n_t & \quad b_t/n_t & \quad a_t b_t/n_t^2 \\
\end{align*}$$

We need $\frac{a_1 b_1}{n_1} + \cdots + \frac{a_t b_t}{n_t} = 0$

Geometrically, we can think of the following partitioning of sums of squares.

\[ SS_{Total} = SS_{C_1} + SS_{C_2} + SS_{Error} = SS_{Treat} + SS_{Error} \]

\[ SS_{C_1} = \text{squared length of orthogonal projection of vector } y - \bar{y}. \text{ onto } C_1 \text{ vector.} \]