7.1 The 1-Way Normal Model

7.1.1 Graphical Comparisons

<table>
<thead>
<tr>
<th>Type 1 Springs</th>
<th>Type 2 Springs</th>
<th>Type 3 Springs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.99, 2.06, 1.99, 1.94, 2.05, 1.88, 2.30</td>
<td>2.85, 2.74, 2.74, 2.63, 2.74, 2.80</td>
<td>2.10, 2.01, 1.93, 2.02, 2.10, 2.05</td>
</tr>
</tbody>
</table>

The data appear
- Fairly normal since straight
- With fairly equal standard deviations since similar slopes, parallel
- One value in the type 1 springs is noticeably different from the others and should be checked.

For all of chapters 7 and 8 we will be assuming:
- Normal populations
  - Straight normal plots
- Equal variances
  - Parallel normal plots
- Independent measurements/ residuals
  - Known from how the experiment was done
  - No repeated measures or blocks
  - Temperature of cans taken over time
    - Correlated measurements/ residuals from same can
7.1.2 The One-Way (Normal) Multisample Model, Fitted Values, Residuals.

<table>
<thead>
<tr>
<th>Groups</th>
<th>n</th>
<th>Mean</th>
<th>Variance</th>
<th>St Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trilene XL</td>
<td>10</td>
<td>11.52</td>
<td>0.01733</td>
<td>0.1317</td>
</tr>
<tr>
<td>Trilene XT</td>
<td>10</td>
<td>11.64</td>
<td>0.01600</td>
<td>0.1265</td>
</tr>
<tr>
<td>Stren</td>
<td>10</td>
<td>11.10</td>
<td>0.01333</td>
<td>0.1155</td>
</tr>
</tbody>
</table>

\[ Y_{ij} = i^{th} \text{ value in } j^{th} \text{ group.} \]
\[ i = 1, \ldots, r. \quad r = \text{number of treatment groups} \]
\[ j = 1, \ldots, n_i. \quad n_i = \text{number of values in } i^{th} \text{ group.} \]

\[ Y_{ij} = \mu_i + \epsilon_{ij} \]
- \( \mu_i \) = true, population mean of \( i^{th} \) treatment
- \( \epsilon_{ij} \) = error
  - Independent normal errors with constant variance \( \sigma^2 \)

Fitted Value in \( i^{th} \) treatment group = Sample mean = \( \bar{Y}_i \)

\[ \hat{Y}_{ij} = \bar{Y}_i \]

Residual = \( e_{ij} = y_{ij} - \hat{Y}_{ij} = y_{ij} - \bar{Y}_i \)

7.1.3 A Pooled Estimate of \( \sigma^2 \)

If all groups have same variance \( \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_r^2 = \sigma^2 \)
- Each group's sample variance, \( S_i^2 \), estimates \( \sigma^2 \).
- Combine these into a pooled estimate of \( \sigma^2 \).

\[ S_p^2 = \frac{\sum (n_i - 1)S_i^2}{\sum (n_i - 1)} \quad \text{Weighted average of sample variances.} \]
\[ \text{Weighted by their } df. \]

\[ df = \sum_{i=1}^{r} (n_i - 1) \quad \text{The pooled } df \text{ is the total of all } df. \]
As with t-tests in Chapter 6, we estimate the variance, $\sigma^2$, with the pooled variance, $S_p^2$.

<table>
<thead>
<tr>
<th>Groups</th>
<th>n</th>
<th>Mean</th>
<th>Variance</th>
<th>St Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trilene XL</td>
<td>10</td>
<td>11.52</td>
<td>0.01733</td>
<td>0.1317</td>
</tr>
<tr>
<td>Trilene XT</td>
<td>10</td>
<td>11.64</td>
<td>0.01600</td>
<td>0.1265</td>
</tr>
<tr>
<td>Stren</td>
<td>10</td>
<td>11.10</td>
<td>0.01333</td>
<td>0.1155</td>
</tr>
</tbody>
</table>

$$S_p^2 = \frac{\sum\sum(n_{ij}-1)S_{ij}^2}{\sum\sum(n_{ij}-1)} = \frac{(10-1)*0.01733 + (10-1)*0.01600 + (10-1)*0.01333}{10-1 + 10-1 + 10-1} = 0.4200 \div 27 = 0.01556$$

$$S_p^2 = \text{MS Error} = \text{MS Residual}$$

$$\frac{\sum\sum(n_{ij}-1)S_{ij}^2}{\sum\sum(n_{ij}-1)} = \frac{\sum\sum(n_{ij}-1)(y_{ijk} - \bar{y}_{ij})^2}{\sum\sum(n_{ij}-1)} = \frac{\text{Sum of Squared Residuals}}{\text{Residual df}} = \text{MS Residual}$$

The Residual MS is called Error MS in the text or Within Groups by Excel.

Excel Anova: Single Factor

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P-value</th>
<th>F crit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>1.6080</td>
<td>2</td>
<td>0.8040</td>
<td>51.69</td>
<td>5.867E-10</td>
<td>3.354</td>
</tr>
<tr>
<td>Within Groups</td>
<td>0.4200</td>
<td>27</td>
<td>0.01556</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2.0280</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$$S_p^2 = 0.01556$$

$$S_p = 0.1247$$

Total SS = $\sum_{i=1}^{r} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{all})^2$
Springs Example  Example 2 on page 452

<table>
<thead>
<tr>
<th>( n_i )</th>
<th>( \bar{Y}_i )</th>
<th>( S_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>2.030</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>2.750</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>2.035</td>
</tr>
</tbody>
</table>

\[
s_p^2 = \frac{6(0.134^2) + 5(0.074^2) + 5(0.064^2)}{6 + 5 + 5} = 0.0097
\]

\[
s_p = \sqrt{0.0097} = 0.099
\]

7.2 Simple Confidence Intervals

\[
\mu_i = \bar{Y}_i \pm t \sqrt{\frac{s_p^2}{n_i}} = \bar{Y}_i \pm t \frac{s_p}{\sqrt{n_i}}
\]

\[
\mu_i - \mu_k = \bar{Y}_i - \bar{Y}_k \pm t \sqrt{\frac{s_p^2}{n_i} + \frac{s_p^2}{n_k}} = \bar{Y}_i - \bar{Y}_k \pm t \cdot s_p \sqrt{\frac{1}{n_i} + \frac{1}{n_k}}
\]

\[
t = \text{tabled } t \text{ with } df = \sum (n_i - 1)
\]

- Chapter 7 uses the pooled variance when finding a confidence interval for \( \bar{Y}_i \).
  - This is different from Chapter 6.
  - If all groups have the same variance, the pooled variance estimate is a better guess at \( \sigma \).
- With small numbers of values in each treatment group, the confidence intervals are more precise.
  - For example with \( n=3 \) values in 6 treatment groups
    - 2 df: \( \bar{Y} \pm 4.303 \cdot SE_F \)
    - Pooled 12 df: \( \bar{Y} \pm 2.179 \cdot SE_F \)
Springs Example

\[ df = (7 - 1) + (6 - 1) + (6 - 1) = 16 \]

<table>
<thead>
<tr>
<th>Spring</th>
<th>( \bar{Y}_i )</th>
<th>( n_i )</th>
<th>( \sqrt{s_p^2/n} )</th>
<th>( 2.12\sqrt{s_p^2/n} )</th>
<th>95% Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.03</td>
<td>7</td>
<td>0.037</td>
<td>0.079</td>
<td>1.95 to 2.11</td>
</tr>
<tr>
<td>2</td>
<td>2.75</td>
<td>6</td>
<td>0.040</td>
<td>0.085</td>
<td>2.66 to 2.83</td>
</tr>
<tr>
<td>3</td>
<td>2.035</td>
<td>6</td>
<td>0.040</td>
<td>0.085</td>
<td>1.95 to 2.12</td>
</tr>
</tbody>
</table>

\[ \bar{Y}_2 - \bar{Y}_1 \pm 2.12 \sqrt{s_p^2 \left( \frac{1}{7} + \frac{1}{6} \right)} \]

0.72 ± 2.12(0.055)
0.72 ± 0.12
0.60 to 0.84  Clearly Different

\[ \bar{Y}_3 - \bar{Y}_1 \pm 2.12(0.055) \]

– 0.11 to 0.12  Similar
### 7.2.2 General Linear Combinations

#### 7.4.1 Significance Testing

\[ L = \sum_{i=1}^{r} c_i \mu_i \] Parameter

\[ \hat{L} = \sum_{i=1}^{r} c_i \bar{Y}_i \] Estimate

\[ \text{Var}(\hat{L}) = \sum \text{Var}(c_i \bar{Y}_i) = \sum_{i=1}^{r} c_i^2 \text{Var}(\bar{Y}_i) = \sum_{i=1}^{r} c_i^2 \frac{\sigma^2}{n_i} = \sigma^2 \sum_{i=1}^{r} c_i^2 \]

\[ \hat{V}(\hat{L}) = s_p^2 \sum_{i=1}^{r} \frac{c_i^2}{n_i} \] Estimated

Confidence Interval

\[ \hat{L} \pm t \frac{S_p}{\sqrt{\sum_{i=1}^{r} \frac{c_i^2}{n_i}}} \quad df = \sum (n_i - 1) \]

**Springs:** Compare the average of spring types 1 and 3 to spring type 2.

\[ \bar{Y}_2 - \frac{\bar{Y}_1 + \bar{Y}_3}{2} = 2.75 - \frac{1}{2}(2.03) - \frac{1}{2}(2.035) \]

\[ 2.75 - \frac{1}{2}(2.03) - \frac{1}{2}(2.035) \pm 2.12 \sqrt{0.0097 \left( \frac{(1/2)^2}{7} + \frac{1^2}{6} + \frac{(1/2)^2}{6} \right)} \]

\[ 0.715 \pm 2.12(0.049) \]

\[ 0.715 \pm 0.103 \]

0.61 to 0.82
# Excel Anova: Single Factor

<table>
<thead>
<tr>
<th>Groups</th>
<th>Count</th>
<th>Sum</th>
<th>Average</th>
<th>Variance</th>
<th>St Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trilene XL</td>
<td>10</td>
<td>115.2</td>
<td>11.52</td>
<td>0.01733</td>
<td>0.1317</td>
</tr>
<tr>
<td>Trilene XT</td>
<td>10</td>
<td>116.4</td>
<td>11.64</td>
<td>0.01600</td>
<td>0.1265</td>
</tr>
<tr>
<td>Stren</td>
<td>10</td>
<td>111.0</td>
<td>11.10</td>
<td>0.01333</td>
<td>0.1155</td>
</tr>
</tbody>
</table>

## ANOVA

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P-value</th>
<th>F crit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>1.6080</td>
<td>2</td>
<td>0.8040</td>
<td>51.69</td>
<td>5.867E-10</td>
<td>3.354</td>
</tr>
<tr>
<td>Within Groups</td>
<td>0.4200</td>
<td>27</td>
<td>0.01556</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2.0280</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
S_p^2 = 0.01556 \\
S_p = 0.1247
\]

From other software

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Standard Error</th>
<th>t</th>
<th>p-value</th>
<th>95% Lower</th>
<th>95% Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>XT - XL</td>
<td>0.120</td>
<td>0.0558</td>
<td>2.15</td>
<td>0.0405</td>
<td>0.006</td>
<td>0.234</td>
</tr>
<tr>
<td>XT - Stren</td>
<td>0.540</td>
<td>0.0558</td>
<td>9.68</td>
<td>&lt; .0001</td>
<td>0.426</td>
<td>0.654</td>
</tr>
<tr>
<td>XL - Stren</td>
<td>0.420</td>
<td>0.0558</td>
<td>7.53</td>
<td>&lt; .0001</td>
<td>0.306</td>
<td>0.534</td>
</tr>
<tr>
<td>(XT+XL)/2 - Stren</td>
<td>0.480</td>
<td>0.0483</td>
<td>9.94</td>
<td>&lt; .0001</td>
<td>0.381</td>
<td>0.579</td>
</tr>
</tbody>
</table>

---

**Line Strength**

<table>
<thead>
<tr>
<th>Type of Fishing Line</th>
<th>0.0</th>
<th>1.0</th>
<th>2.0</th>
<th>3.0</th>
<th>4.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>1=Trilene XL</td>
<td>110</td>
<td>112</td>
<td>114</td>
<td>116</td>
<td>118</td>
</tr>
<tr>
<td>2=Trilene XT</td>
<td>115</td>
<td>117</td>
<td>119</td>
<td>120</td>
<td>121</td>
</tr>
<tr>
<td>3=Stren</td>
<td>116</td>
<td>118</td>
<td>120</td>
<td>122</td>
<td>124</td>
</tr>
<tr>
<td>4=All</td>
<td>119</td>
<td>121</td>
<td>123</td>
<td>125</td>
<td>127</td>
</tr>
</tbody>
</table>
Example 5 on page 467

<table>
<thead>
<tr>
<th>Heat Treat</th>
<th>% water</th>
<th>Brick Strength</th>
<th>$\bar{x}$</th>
<th>$s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slow</td>
<td>17</td>
<td>4911, 5958, 5676</td>
<td>5528</td>
<td>558</td>
</tr>
<tr>
<td>Slow</td>
<td>19</td>
<td>4387, 5388, 5007</td>
<td>4927</td>
<td>505</td>
</tr>
<tr>
<td>Fast</td>
<td>17</td>
<td>3824, 3140, 3502</td>
<td>3489</td>
<td>312</td>
</tr>
<tr>
<td>Fast</td>
<td>19</td>
<td>4768, 3672, 3242</td>
<td>3894</td>
<td>787</td>
</tr>
</tbody>
</table>

First Step: PLOT THE DATA

Heat effect

$$\bar{Y}_{slow} = \frac{1}{2}(5528) + \frac{1}{2}(4927) = 5227.5$$

$$\bar{Y}_{fast} = \frac{1}{2}(3489) + \frac{1}{2}(3894) = 3691.5$$

With equal n's

$$\bar{Y}_{slow} - \bar{Y}_{fast} = \frac{1}{2}Y_{11} + \frac{1}{2}Y_{12} - \frac{1}{2}Y_{21} - \frac{1}{2}Y_{22}$$

$$\hat{L} = \frac{1}{2}Y_{11} + \frac{1}{2}Y_{12} - \frac{1}{2}Y_{21} - \frac{1}{2}Y_{22} = 1536$$

$$\sqrt{\hat{V}(\hat{L})} = S_p \sqrt{\frac{(1/2)^2 \cdot 4}{3}} = S_p \sqrt{\frac{1}{3}}$$

$$= 570.7 \cdot 0.577$$

$$= 329.6$$
CI 1536 ± 2.306(329.6)  
1536 ± 760  
776 to 2296

With equal numbers of values in each group, the comparison of the heat levels is equivalent to

\[
\bar{Y}_{\text{slow}} = \text{average of 6 values} \\
\bar{Y}_{\text{fast}} = \text{average of 6 values} \\
\sqrt{\hat{V}(\bar{Y}_{\text{slow}} - \bar{Y}_{\text{fast}})} = S_p \sqrt{\frac{1}{6} + \frac{1}{6}}
\]

The same as above with

\[
\sqrt{\hat{V}(\hat{L})} = S_p \sqrt{\frac{(1/2)^2 \ast 4}{3}}
\]

To test \( H_0 \): Heat main effect = 0  
\( H_a \): Heat main effect ≠ 0

\[
t = \frac{1536}{329.6} = 4.66
\]

\( 2 \ast 0.005 < p-value < 2 \ast 0.001 \)
\( 0.001 < p-value < 0.002 \)

Water main effect

\[
\hat{L} = \frac{1}{2} (5528 + 3489) - \frac{1}{2} (4927 + 3894) = 98
\]

\[
\sqrt{\hat{V}(\hat{L})} = 329.6 \quad \text{(same as heat effect)}
\]

Obviously, there is not much evidence of a water main effect.