A *posteriori* multiple comparison tests
Recall the Lakes “experiment”

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lakes</td>
<td>58.000</td>
<td>2</td>
<td>29.400</td>
<td>8.243</td>
<td>0.006</td>
</tr>
<tr>
<td>Error</td>
<td>42.800</td>
<td>12</td>
<td>3.567</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>101.600</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- The ANOVA tells us that there are differences in average fish length among levels of factor Lakes \( (p<0.05) \).
- Factor Lakes has three levels (Lake 1, Lake 2 and Lake 3)
- Which Lakes are different? Note that there are several possibilities! Are they all different? Is Lake 1 different from Lake 2 which, in turn, is equal to Lake 3?
Recall the Lakes “experiment”

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lakes</td>
<td>58.000</td>
<td>2</td>
<td>29.400</td>
<td>8.243</td>
<td>0.006</td>
</tr>
<tr>
<td>Error</td>
<td>42.800</td>
<td>12</td>
<td>3.567</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>101.600</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- The ANOVA avoids Type I Error inflation when analyzing simultaneously more than two samples. However, in its simpler form, it does not give any detail on particular differences among levels.

- However, we can be confident (with an error rate of 0.05) that some of the levels are really different.
A posteriori Multiple Comparison Tests

- These type of tests is used when we have no a priori knowledge of the type of differences that are expected.
- They work like multiple t-tests but take advantage of the results of ANOVA (e.g., the estimated Error or Within Samples variation).
- They are prone to Type I error inflation, but try to avoid that by specific strategies (algorithms).
- There is a plethora of tests to choose from!
A posteriori Multiple Comparison Tests

- Student-Neuman-Keuls test (SNK)
- Tukey Honestly Significant Difference (HSD or Tukey)
- Fisher Protected Least Significant Difference (LSD)
- Duncan Multiple Range test (Duncan)
- Ryan’s test
- Peritz’s test
- Scheffé’s test
- Dunnett’s test
- Sequential Bonferroni Correction
- Dunnet’s T3 test (rank based)
- Dunnet’s C test (rank based)
- Games-Howell test (rank based)
- ...
Tukey Honestly Significant Difference (HSD or Tukey)

\[ Q = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\frac{MS_{\text{error}}}{n}}} \]

The basic formula is very simple, similar to a t-test, but uses \( MS_{\text{error}} \) instead of the variances of the samples.
Experimental Design

Multiple Tests

Tukey Honestly Significant Difference (HSD or Tukey)

1 - First of all, compute the averages for each sample

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Average</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lake 1</td>
<td>5</td>
<td>13.2</td>
<td>1.7</td>
</tr>
<tr>
<td>Lake 2</td>
<td>5</td>
<td>15.0</td>
<td>3.5</td>
</tr>
<tr>
<td>Lake 3</td>
<td>5</td>
<td>18.0</td>
<td>5.5</td>
</tr>
</tbody>
</table>
Tukey Honestly Significant Difference (HSD or Tukey)

2 – Order the averages in ascending or descending order (it doesn’t matter)

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Average</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lake 1</td>
<td>5</td>
<td>13.2</td>
<td>1.7</td>
</tr>
<tr>
<td>Lake 2</td>
<td>5</td>
<td>15.0</td>
<td>3.5</td>
</tr>
<tr>
<td>Lake 3</td>
<td>5</td>
<td>18.0</td>
<td>5.5</td>
</tr>
</tbody>
</table>

Lake1 13.2 Lake2 15.0 Lake3 18.0
Tukey Honestly Significant Difference (HSD or Tukey)

3 – Get the $\text{MS}_{\text{error}}$, $\text{DF}_{\text{error}}$, and $n$ from the ANOVA

Lake1  Lake2  Lake3
13.2    15.0    18.0

$\text{MS}_{\text{Error}} = 3.567$
$\text{DF}_{\text{Error}} = 12$
$n = 5$
Tukey Honestly Significant Difference (HSD or Tukey)

4 – Compare the two most extreme averages using the formula of $Q$

\[
Q = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\frac{MS_{\text{Error}}}{n}}} = \frac{|13.2 - 18.0|}{\sqrt{\frac{3.567}{5}}} = 5.68296
\]

$Q_{\text{obs}} = 5.6829$

Lake1: 13.2
Lake2: 15.0
Lake3: 18.0

$\text{DF}_{\text{Error}} = 12$
$n = 5$

$\text{MS}_{\text{Error}} = 3.567$
Tukey Honestly Significant Difference (HSD or Tukey)

4 – Obtain a critical value for the distribution of $Q$ with $k=3$ means and $\nu = DF_{\text{error}} = 12$ degrees of freedom

$$Q_{\text{obs}} = 5.6829 \quad q_{[3,12]} = 3.773$$

5 – If the observed $Q$ is larger than $q_{\text{crit}}$, then the two means differ

Lake 1 ≠ Lake 3
Tukey Honestly Significant Difference (HSD or Tukey)

6 – Now repeat these steps for Lake 1 vs. Lake 2 and Lake 2 vs. Lake 3, simultaneously maintaining $q_{crit}$

For Lake 2 vs. Lake 3

\[
Q = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\frac{MS_{Error}}{n}}} = \frac{|15.0 - 18.0|}{\sqrt{\frac{3.567}{5}}} = 3.5518
\]

$MS_{Error} = 3.567$

$DF_{Error} = 12$

$n = 5$

$Q_{obs} < q_{crit}$

Lake 2 = Lake 3
Experimental Design

Multiple Tests

Tukey Honestly Significant Difference (HSD or Tukey)

6 – Now repeat these steps for Lake 1 vs. Lake 2 and Lake 2 vs. Lake 3, simultaneously maintaining $q_{crit}$

For Lake 1 vs. Lake 2

<table>
<thead>
<tr>
<th>Lake1</th>
<th>Lake2</th>
<th>Lake3</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.2</td>
<td>15.0</td>
<td>18.0</td>
</tr>
</tbody>
</table>

$Q = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\frac{MS_{Error}}{n}}} = \frac{|13.2 - 15.0|}{\sqrt{\frac{3.567}{5}}} = 2.2495$

$Q_{obs} < q_{crit}$

Lake 1 = Lake 2

$MS_{Error} = 4.167$

$DF_{Error} = 12$

$n = 5$

$q_{[3,12]} = 3.773$
Experimental Design

Multiple Tests

Tukey Honestly Significant Difference (HSD or Tukey)

Lake1  | Lake2 | Lake3  
13.2   | 15.0  | 18.0   

The multiple test is inconclusive!

The most likely reason for such outcome is the lack of power.

You should repeat the analysis with larger sample sizes ($n$)
Student-Neuman-Keuls test (SNK)

1 - First of all, compute the averages for each sample

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Average</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lake 1</td>
<td>5</td>
<td>13.2</td>
<td>1.7</td>
</tr>
<tr>
<td>Lake 2</td>
<td>5</td>
<td>15.0</td>
<td>3.5</td>
</tr>
<tr>
<td>Lake 3</td>
<td>5</td>
<td>18.0</td>
<td>5.5</td>
</tr>
</tbody>
</table>
Tukey Honestly Significant Difference (HSD or Tukey)

2 – Order the averages in ascending or descending order (it doesn’t matter)

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Average</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lake 1</td>
<td>5</td>
<td>13.2</td>
<td>1.7</td>
</tr>
<tr>
<td>Lake 2</td>
<td>5</td>
<td>15.0</td>
<td>3.5</td>
</tr>
<tr>
<td>Lake 3</td>
<td>5</td>
<td>18.0</td>
<td>5.5</td>
</tr>
</tbody>
</table>

Lake1 13.2 Lake2 15.0 Lake3 18.0
Experimental Design

Student-Neuman-Keuls test (SNK)

3 – Get the $MS_{\text{error}}$, $DF_{\text{error}}$ and $n$ from the ANOVA

Lake1 13.2  
Lake2 15.0  
Lake3 18.0  

$MS_{\text{Error}} = 3.567$

$DF_{\text{Error}} = 12$

$n = 5$
Experimental Design

Multiple Tests

Student-Neuman-Keuls test (SNK)

4 – Compare the two most extreme averages using the formula of Q

\[ Q_{\text{obs}} = 5.6829 \]

Lake1  13.2  Lake2  15.0  Lake3  18.0

\[ Q = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\frac{\text{MS}_{\text{Error}}}{n}}} = \frac{|13.2 - 18.0|}{\sqrt{\frac{3.567}{5}}} = 5.68296 \]

\[ \text{MS}_{\text{Error}} = 3.567 \]

\[ \text{DF}_{\text{Error}} = 12 \]

\[ n = 5 \]

\[ Q_{\text{obs}} = 5.6829 \]
Experimental Design

Multiple Tests

Student-Neuman-Keuls test (SNK)

4 – Obtain a critical value for the distribution of Q with k=3 means and $\nu = \text{DF}_{\text{error}} = 12$ degrees of freedom

$$Q_{\text{obs}} = 5.6829$$  $$q_{[3,12]} = 3.773$$

5 – If the observed Q is larger than $q_{\text{crit}}$, then the two means differ

Lake 1 $\neq$ Lake 3
Experimental Design

Multiple Tests

**Student-Neuman-Keuls test (SNK)**

6 – Now repeat these steps for Lake 1 vs. Lake 2 and Lake 2 vs. Lake 3, simultaneously, **but use a** $q_{\text{crit}}$ **for two means instead of three**

<table>
<thead>
<tr>
<th>Lake 1</th>
<th>Lake 2</th>
<th>Lake 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.2</td>
<td>15.0</td>
<td>18.0</td>
</tr>
</tbody>
</table>

For Lake 2 vs. Lake 3

\[
Q = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\frac{\text{MS}_{\text{Error}}}{n}}} = \frac{|15.0 - 18.0|}{\sqrt{\frac{3.567}{5}}} = 3.5518
\]

\[
q = 3.082
\]

\[
Q_{\text{obs}} > q_{\text{crit}} \quad \text{Lake 2} \neq \text{Lake 3}
\]

**Q** $\text{Error} = 3.567$

**DF** $\text{Error} = 12$

**n** = 5

**$q_{[2,12]}$** = 3.082
Experimental Design

Multiple Tests

Student-Neuman-Keuls test (SNK)

6 – Now repeat these steps for Lake 1 vs. Lake 2 and Lake 2 vs. Lake 3, simultaneously, but use a $q_{crit}$ for two means instead of three

For Lake 1 vs. Lake 2

<table>
<thead>
<tr>
<th>Lake1</th>
<th>Lake2</th>
<th>Lake3</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.2</td>
<td>15.0</td>
<td>18.0</td>
</tr>
</tbody>
</table>

$$Q = \frac{|\bar{X}_1 - \bar{X}_2|}{\frac{\text{MS}_{\text{Error}}}{n}} = \frac{13.2 - 15.0}{\sqrt{3.567/5}} = 2.2495$$

$Q_{obs} < q_{crit}$

Lake 1 = Lake 2

$MS_{\text{Error}} = 3.567$

$DF_{\text{Error}} = 12$

$n = 5$

$q_{[2,12]} = 3.082$
The multiple test is conclusive!

There are two homogeneous groups:
Lake 1 and Lake 2 (they do not differ)
Lake 3 (which differs from the group above)

<table>
<thead>
<tr>
<th></th>
<th>Lake1</th>
<th>Lake2</th>
<th>Lake3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>13.2</td>
<td>15.0</td>
<td>18.0</td>
</tr>
</tbody>
</table>

Student-Neuman-Keuls test (SNK)