data bal;
input y group;
cards;
0 1
0 1
4 1
4 1
2 2
2 2
3 2
3 2
4 2
4 2
0 3
1 3
2 3
2 3
3 3
3 3
4 4
4 4
5 4
5 4
7 4
9 4
The GLM Procedure
Class Level Information
Class         Levels    Values
            4    1 2 3 4
Number of Observations Used          24
Level of        ----------------------y------------------
            group        N             Mean          Std Dev
            1            4       2.00000000       2.30940108
            2            8       3.00000000       0.75592895
            3            7       2.00000000       1.15470054
            4            5       6.00000000       2.00000000
The GLM Procedure
Dependent Variable: y
Source                      DF         Squares     Mean Square    F Value    Pr > F
Model                        3     55.33333333     18.44444444       8.38    0.0008
Error                       20     44.00000000      2.20000000
Corrected Total             23     99.33333333
R-Square     Coeff Var      Root MSE        y Mean
0.557047      46.83915      1.483240      3.166667
O'Brien's Test for Homogeneity of y Variance
ANOVA of O'Brien's Spread Variable, W = 0.5
Source        DF     Squares     Mean    F Value    Pr > F
group          3     81.6992     27.2331       4.37    0.0161
Error         20     124.7      6.2366
Levene's Test for Homogeneity of y Variance
ANOVA of Squared Deviations from Group Means
Source        DF     Squares     Mean    F Value    Pr > F
group          3     45.6762     15.2254       4.94    0.0100
Error         20     61.6571      3.0829
Welch's ANOVA for y
Source        DF     F Value    Pr > F
            group          3.0000     4.84      0.0344
            Error          7.7786
Beasley (BST 622) Post-Hoc Tests and Planned Comparisons
**t Tests (LSD) for y**

NOTE: This test controls the Type I comparisonwise error rate, not the experimentwise error rate.

<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>2.2</td>
</tr>
<tr>
<td>Critical Value of t</td>
<td><strong>2.08596</strong></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Difference</th>
<th>95% Confidence Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - 2</td>
<td>-1.0000</td>
<td>-2.8947   0.8947</td>
</tr>
<tr>
<td>1 - 3</td>
<td>0.0000</td>
<td>-1.9393   1.9393</td>
</tr>
<tr>
<td>1 - 4</td>
<td>-4.0000</td>
<td>-6.0755   -1.9245 ***</td>
</tr>
<tr>
<td>2 - 3</td>
<td>1.0000</td>
<td>-0.6013   2.6013</td>
</tr>
<tr>
<td>2 - 4</td>
<td>-3.0000</td>
<td>-4.7638   -1.2362 ***</td>
</tr>
<tr>
<td>3 - 4</td>
<td>-4.0000</td>
<td>-5.8117   -2.1883 ***</td>
</tr>
</tbody>
</table>

The Fisher’s LSD is simply a *t*-test with the variances pooled over groups as the error term. The *df* is simply *df* for the model; *df* = N-J. There is no correction for multiple testing.

The LSD is defined as: CV x SE

The critical value (CV) = \( t_{(1-\alpha/2),(N-J)} \). With \( \alpha = 0.05 \), \( N = 24 \), \( J = 4 \) groups, \( t_{(0.975,(20))} = **2.08596** \).

The Standard Error (SE) is uses the Root Mean Square Error RMSE = \( \sqrt{2.2} = **1.483240** \)

\[
SE_{(LSD)} = \sqrt{MSE\left[\frac{1}{n_j} + \frac{1}{n_{j'}}\right]}
\]

Thus, LSD = \( t_{(1-\alpha/2),(N-J)} SE_{(LSD)} = t_{(1-\alpha/2),(N-J)} \sqrt{MSE\left[\frac{1}{n_j} + \frac{1}{n_{j'}}\right]} \)

100(1-\( \alpha \))% Confidence Intervals are constructed as

\[
(\bar{Y}_j - \bar{Y}_{j'}) \pm t_{(1-\alpha/2),(N-J)} SE_{(LSD)}
\]

The test statistic is formed as \( t_{(LSD,df=N-J)} = \frac{\bar{Y}_j - \bar{Y}_{j'}}{\sqrt{MSE\left[\frac{1}{n_j} + \frac{1}{n_{j'}}\right]}} \)

For the Comparison of Group 3 and 4

\[
LSD = **2.08596** \times \sqrt{2.2\left[\frac{1}{7} + \frac{1}{5}\right]} = **2.08596** \times **0.8685** = **1.81165**.
\]

The 95% Confidence Interval is: -4 ± 1.8117  [-5.8117  -2.1883]

The LSD t-test is \( t_{(df=20)} = -4/0.8685 = -4.606 \), which exceeds the critical value of \( t_{(0.975,(20))} = **2.08596** \). The 2-tailed p-value for \( t_{(df=20)} = -4.606 \), is \( p = 0.0001711 \), which can be found with the SAS code: pt=2 * (1-probt (4.605662, 20));
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 Tukey’s for all pairwise comparisons.

<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>2.2</td>
</tr>
<tr>
<td>Critical Value of t</td>
<td>2.92712</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>1 - 2</td>
<td>-1.0000</td>
<td>-3.6587  1.6587</td>
</tr>
<tr>
<td>1 - 3</td>
<td>0.0000</td>
<td>-2.7213  2.7213</td>
</tr>
<tr>
<td>1 - 4</td>
<td>-4.0000</td>
<td>-6.9124  -1.0876 ***</td>
</tr>
<tr>
<td>2 - 3</td>
<td>1.0000</td>
<td>-1.2470  3.2470</td>
</tr>
<tr>
<td>2 - 4</td>
<td>-3.0000</td>
<td>-5.4751  -0.5249 ***</td>
</tr>
<tr>
<td>3 - 4</td>
<td>-4.0000</td>
<td>-6.5422  -1.4578 ***</td>
</tr>
</tbody>
</table>

The Bonferroni correction of Fisher’s LSD is simply a \( t \)-test with the variances pooled over groups as the error term. The \( df \) is simply \( df \) for the model; \( df = N\text{-}J \). The alpha is adjusted by dividing alpha by the number of pairwise comparisons.

\[ \alpha_{BON} = \alpha / [J(J-1)/2] \]

The Bonferroni adjusted LSD is defined as: \( CV \times SE \)

The critical value (CV) = \( t_{[(1-\alpha(BON)/2),(N\text{-}J)]} \). With \( \alpha = 0.05 \), \( N = 24 \), \( J = 4 \) groups,

\[ \alpha_{BON} = \alpha / 6 = 0.05/6 = 0.00833 \text{ and } t_{[0.99583,(20)]} = 2.92712. \]

Thus, \( LSD_{(BON)} = t_{[(1-\alpha/2),(N\text{-}J)]} \times MSE_{(LSD)} \) = \( t_{[(1-\alpha(BON)/2),(N\text{-}J)]} \sqrt{MSE \left( \frac{1}{n_j} + \frac{1}{n_{j'}} \right)} \)

100(1-\( \alpha \))% Simultaneous Confidence Intervals are constructed as

\[ (\bar{Y}_j - \bar{Y}_{j'}) \pm t_{[(1-\alpha(BON)/2),(N\text{-}J)]} \times MSE_{(LSD)} \]

The test statistic is formed as \( t_{(BON, df = N\text{-}J)} = \frac{\bar{Y}_j - \bar{Y}_{j'}}{\sqrt{MSE \left( \frac{1}{n_j} + \frac{1}{n_{j'}} \right)}} \)

For the Comparison of Group 3 and 4

\[ LSD_{(BON)} = 2.92712 \times \sqrt{2.2\left[\frac{1}{7} + \frac{1}{5}\right]} = 2.92712 \times 0.8685 = 2.5422. \]

The 95% Confidence Interval is: -4 ± 2.5422 [-6.5422 -1.4578]

The t-test is \( t_{(df = 20)} = -4/0.8685 = -4.606 \), which exceeds the critical value of \( t_{[0.975,(20)]} = 2.92712 \). A Bonferroni adjusted p-value is created by reversing the Bonferroni correction and multiply the p-value by the number of tests \( p_{(BON)} = 0.0001711 \times 6 = 0.0010266 \)
Sidak 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 Tukey’s for all pairwise comparisons.

<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>2.2</td>
</tr>
<tr>
<td>Critical Value of t</td>
<td>2.91761</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Difference</th>
<th>group</th>
<th>Between Means</th>
<th>Simultaneous 95% Confidence Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - 2</td>
<td>-1.0000</td>
<td>-3.6501</td>
<td>1.6501</td>
</tr>
<tr>
<td>1 - 3</td>
<td>0.0000</td>
<td>-2.7124</td>
<td>2.7124</td>
</tr>
<tr>
<td>1 - 4</td>
<td>-4.0000</td>
<td>-6.9030</td>
<td>-1.0970 ***</td>
</tr>
<tr>
<td>2 - 3</td>
<td>1.0000</td>
<td>-1.2397</td>
<td>3.2397</td>
</tr>
<tr>
<td>2 - 4</td>
<td>-3.0000</td>
<td>-5.4671</td>
<td>-0.5329 ***</td>
</tr>
<tr>
<td>3 - 4</td>
<td>-4.0000</td>
<td>-6.5339</td>
<td>-1.4661 ***</td>
</tr>
</tbody>
</table>

The Šidák correction of Fisher’s LSD is simply a $t$-test with the variances pooled over groups as the error term. The $df$ is simply $df$ for the model; $df = N-J$. The alpha is adjusted by the number of pairwise comparisons.

$$\alpha_{\text{Šidák}} = 1-[1-\alpha]^{1/k}$$  where $k = [J(J-1)/2]$

The Šidák adjusted LSD is defined as: CVxSE

The critical value (CV) $= t_{[(1-\alpha(\text{Šidák})/2),(N-J)]}$. With $\alpha = 0.05$, $N = 24$, $J = 4$ groups,

$$\alpha_{\text{Šidák}} = 1-[0.95]^{1/6} = 0.008512445$$

and $t_{[0.995743778,(20)]} = 2.91761$.

Thus, $\text{LSD}_{\text{(Šidák)}} = t_{[(1-\alpha/2),(N-J)]}\sqrt{\text{SE}_{(LSD)}} = t_{[(1-\alpha(\text{Šidák})/2),(N-J)]}\sqrt{\text{MSE} \left[ \frac{1}{n_j} + \frac{1}{n_{j'}} \right]}$

100(1-$\alpha$)% Simultaneous Confidence Intervals are constructed as

$$(\overline{Y}_j - \overline{Y}_{j'}) \pm t_{[(1-\alpha(\text{Šidák})/2),(N-J)]}\sqrt{\text{MSE} \left[ \frac{1}{n_j} + \frac{1}{n_{j'}} \right]}$$

The test statistic is formed as

$$t_{(\text{Šidák},df=N-J)} = \frac{\overline{Y}_j - \overline{Y}_{j'}}{\sqrt{\text{MSE} \left[ \frac{1}{n_j} + \frac{1}{n_{j'}} \right]}}$$

For the Comparison of Group 3 and 4

$$\text{LSD}_{\text{(Šidák)}} = 2.91761 \times \sqrt{2.2\left[\frac{1}{7} + \frac{1}{5}\right]} = 2.91761 \times 0.8685 = 2.5339.$$

The 95% Confidence Interval is: $-4 \pm 2.5339 [\text{-6.5339 to -1.4661}]$

The t-test is $t_{(df=20)} = -4/0.8685 = -4.606$, which exceeds the critical value of $t_{[0.995743778,(20)]} = 2.91761$. A Šidák adjusted $p$-value is created by reversing the Šidák correction,

$$p_{\text{(Šidák)}} = 1-[1-p]^6 = 1-(1-0.0001711)^6 = 0.001026161.$$
From SPSS. Dependent Variable: y Tamhane

<table>
<thead>
<tr>
<th>(I) group</th>
<th>(J) group</th>
<th>Difference (I-J)</th>
<th>Mean</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>-1.000</td>
<td>1.185</td>
<td>.974</td>
<td>.001</td>
<td>-7.632</td>
<td>-1.185</td>
<td>5.632</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>0.000</td>
<td>1.234</td>
<td>1.000</td>
<td>.185</td>
<td>-6.087</td>
<td>6.087</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>-4.000</td>
<td>1.461</td>
<td>.974</td>
<td>.001</td>
<td>-9.597</td>
<td>1.597</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1.000</td>
<td>.512</td>
<td>.389</td>
<td>.001</td>
<td>-6.666</td>
<td>2.666</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>-3.000</td>
<td>.934</td>
<td>.144</td>
<td>.001</td>
<td>-7.044</td>
<td>1.044</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>-4.000(*)</td>
<td>.995</td>
<td>.042</td>
<td>.001</td>
<td>-7.852</td>
<td>-1.148</td>
<td></td>
</tr>
</tbody>
</table>

* The mean difference is significant at the .05 level.

The Tamhane T2 modification of Fisher’s LSD uses unpooled variances.

The Standard Error (SE) is: \( SE(T2) = \sqrt{\frac{s^2_j + s^2_{j'}}{n_j}} \). For the Comparison of Group 3 and 4

\[
SE(T2) = \sqrt{\frac{s^2_j + s^2_{j'}}{n_j}} = \sqrt{\frac{1.1547^2}{7} + \frac{2^2}{5}} = \sqrt{0.1905 + .8} = 0.995
\]

The Welch-Satterthwaite correction for the \( df \) is

\[
v = \frac{\left[ \frac{s^2_j + s^2_{j'}}{n_j(n_j-1)} \right]^2}{\frac{s^4_j}{n^2_j(n_j-1)} + \frac{s^4_{j'}}{n^2_{j'}(n_j'-1)}}
\]

Comparing Groups 3 and 4, \( v = \frac{\left[ \frac{1.1547^2}{7} + \frac{2^2}{5} \right]^2}{\frac{1.1547^4}{7^2(6)} + \frac{2^4}{5^2(4)}} = \frac{.9905^2}{1.7777 + \frac{16}{100}} = \frac{.98109}{0.006047 + .16} = 5.908.

The alpha is adjusted by the number of pairwise comparisons using the Šidák correction.

\[
\alpha_{\text{Šidák}} = 1-\left[1-\alpha\right]^{1/k}, \quad \text{where } k = \left[J(J-1)/2\right].
\]

The Šidák adjusted LSD is defined as: \( CV \times SE \)

The critical value (CV) = \( t_{\left[1-(1-\alpha_{\text{Šidák}})/2,v\right]} \). With \( \alpha = 0.05 \), \( N = 24 \), \( J = 4 \) groups, \( \alpha_{\text{Šidák}} = 1-\left[0.95\right]^{1/4} = 0.008512445 \) and \( t_{\left[0.995743778,(5.908)\right]} = 3.8705. \)

100(1-\(\alpha\))% Simultaneous Confidence Intervals are constructed as

\[
(Y_j - \bar{Y}_{j'}) \pm t_{\left[1-(1-\alpha_{\text{Šidák}})/2,v\right]} SE(T2)
\]

The 95% Simultaneous Confidence Interval is: -4 ± (0.995)(3.8705) [-7.852 -1.148]

**Note:** The CV can be back calculated as \( CV = \frac{\text{ABS}[UL-LB]/2SE}{-7.852} = \frac{-7.852}{-1.148/2} = 7.704/1.99 \approx 3.71 \)

The SAS function: \( cv = \text{tinv}(0.995743778,5.908) \) can be used as well.

The test statistic is formed as \( t_{(df=5.908)} = \frac{\bar{Y}_j - \bar{Y}_{j'}}{\sqrt{\frac{s^2_j}{n_j} + \frac{s^2_{j'}}{n_{j'}}}} \)

The t-test is \( t_{(df=5.908)} = -4/0.995 = -4.019, \) which exceeds the critical value of \( t_{(0.995743778,(5.908))} = 3.8705. \) The 2-tailed p-value for \( t_{(df=5.908)} = -4.019, \) is \( p = 0.007192, \) which can be found with the SAS code: \( pt = 2*(1 - \text{probt(4.605662,5.908)}) \); A Šidák adjustment is then made, \( p_{(T2)} = 1-(1-0.007192)^6 = 0.042383527. \) SAS Code = \( pt2 = 1-(1-(2*(1-\text{probt(4.019,5.908)})))^*6; \)
Tukey’s Studentized Range (HSD) Test for y

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

<table>
<thead>
<tr>
<th>Difference</th>
<th>group Comparison</th>
<th>Between Means</th>
<th>Simultaneous 95% Confidence Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - 2</td>
<td>-1.0000</td>
<td>-3.5423</td>
<td>1.5423</td>
</tr>
<tr>
<td>1 - 3</td>
<td>0.0000</td>
<td>-2.6021</td>
<td>2.6021</td>
</tr>
</tbody>
</table>
| 1 - 4               | -4.0000          | -6.7849       | -1.2151                           **
| 2 - 3               | 1.0000           | -1.1486       | 3.1486                            |
| 2 - 4               | -3.0000          | -5.3667       | -0.6333                           **
| 3 - 4               | -4.0000          | -6.4309       | -1.5691                           **

The Tukey and Tukey-Kramer HSD uses the variances pooled over groups as the error term. The \( df \) is simply \( df \) for the model; \( df = N-J \).

The HSD is defined as: \( CV \times SE \). The critical value (CV) is derived from the Studentized Range \( q \) Distribution \( q[(1-\alpha),(v),(J)] \). With \( \alpha = 0.05, N = 24, J = 4 \) groups, \( q[(0.95),(20),(4)] = 3.95829 \).

**Note:** This can be found with the SAS code: \( CVq = \text{PROBMC}("RANGE", . . . 95, 20, 4) \);

The Standard Error (SE) is uses the Root Mean Square Error \( RMSE = \sqrt{2.2} = 1.483240 \)

\[
SE_{(HSD)} = \sqrt{\frac{MSE}{2} \times \left[ \frac{1}{n_j} + \frac{1}{n_j'} \right]}
\]

Thus, \( HSD = q_{[(1-\alpha),(v),(J)]} \times SE_{(HSD)} = q_{[(1-\alpha),(v),(J)]} \sqrt{\frac{MSE}{2} \times \left[ \frac{1}{n_j} + \frac{1}{n_j'} \right]} \)

100(1-\( \alpha \))% Confidence Intervals are constructed as

\[
(\bar{Y}_j - \bar{Y}_{j'}) \pm q_{[(1-\alpha),(N-J)]} \times SE_{(HSD)}
\]

The test statistic is formed as

\[
t_{(HSD, df=N-J)} = \frac{\bar{Y}_j - \bar{Y}_{j'}}{\sqrt{\frac{MSE}{2} \times \left[ \frac{1}{n_j} + \frac{1}{n_j'} \right]}}
\]

For the Comparison of Group 3 and 4

\[
HSD = 3.95829 \times \sqrt{\frac{2.2}{2} \times \left[ \frac{1}{7} + \frac{1}{5} \right]} = 3.95829 \times 0.61412 = 2.4309.
\]

The 95% Confidence Interval is: \(-4 \pm 2.4309 [-6.4309 \ -1.5691]\)

The HSD \( t \)-test is \( t_{(df=20)} = -4/0.61412 = -6.5134 \), which exceeds the critical value of \( q_{[(0.95),(20),(4)]} = 3.95829 \). The 2-tailed \( p \)-value for the Tukey-Kramer \( t_{(df=20)} = -6.5134 \), is \( p = 0.000908 \), which can be found with the SAS code:

\[
\text{probq}=1-\text{PROBMC}("RANGE", 6.5134, . . . 20, 4);
\]

Beasley (BST 622) Post-Hoc Tests and Planned Comparisons
From SPSS. Dependent Variable: y Games-Howell

<table>
<thead>
<tr>
<th></th>
<th>(I) group</th>
<th>(J) group</th>
<th>Difference (I-J)</th>
<th>Error</th>
<th>Sig.</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>-1.000</td>
<td>1.185</td>
<td>.833</td>
<td></td>
<td>-6.346</td>
<td>4.346</td>
</tr>
<tr>
<td>3</td>
<td>.000</td>
<td>-4.000</td>
<td>1.234</td>
<td>1.000</td>
<td></td>
<td>-5.104</td>
<td>5.104</td>
</tr>
<tr>
<td>4</td>
<td>-3.000</td>
<td>1.000</td>
<td>1.461</td>
<td>.117</td>
<td></td>
<td>-9.044</td>
<td>1.044</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>-4.000(*)</td>
<td>.995</td>
<td>.027</td>
<td></td>
<td>-7.462</td>
<td>-.538</td>
</tr>
</tbody>
</table>

* The mean difference is significant at the .05 level.

The Games-Howell modification of the Tukey-Kramer HSD uses unpooled variances.

The Standard Error (SE) is: 
\[ SE_{(GH)} = \sqrt{\frac{s_j^2 + s_{j'}^2}{n_j + n_{j'}}/2} \]

For the Comparison of Group 3 and 4
\[ SE_{(GH)} = \sqrt{\frac{1.15472^2 + 2^2}{7}}/2 = \sqrt{0.9905}/2 = 0.70374 \]

The Welch-Satterthwaite correction for the df is 
\[ v = \frac{\left(\frac{1.15472}{7} + \frac{2^2}{5}\right)^2}{\left(\frac{1}{1.7777} + \frac{16}{100}\right)} \]

Comparing Groups 3 and 4, 
\[ v = \frac{\left(\frac{1.15474}{7^2(6)} + \frac{2^4}{5^2(4)}\right)^2}{\left(\frac{1.7777}{294} + \frac{16}{100}\right)} = \frac{.9905^2}{0.006047 + .16} = .98109 = 5.908. \]

Thus, the critical value (CV) from the Studentized Range (q) Distribution \[ q_{[(1-\alpha),(v),(J)]} \] With \( \alpha = 0.05 \), \( v = 5.908 \), \( J = 4 \) groups, \( q_{[(0.95),(5.908),(4)]} = 4.919750 \).

Note: This can be found with the SAS code: \[ \text{CVq} = \text{PROBMC}("RANGE", . . . , 95, 5.908, 4); \]

Thus, \( \text{GH-HSD} = q_{[(1-\alpha),(v),(J)]} SE_{(GH)} = q_{[(1-\alpha),(v),(J)]} \sqrt{\frac{s_j^2 + s_{j'}^2}{n_j + n_{j'}}/2} \)

100(1-\( \alpha \))% Confidence Intervals are constructed as 
\[ (\bar{Y}_j - \bar{Y}_{j'}) \pm q_{[(1-\alpha),(v),(J)]} SE_{(GH)} \]

The test statistic is formed as 
\[ t_{(GH,df=5)} = \frac{\bar{Y}_j - \bar{Y}_{j'}}{\sqrt{\frac{s_j^2 + s_{j'}^2}{n_j + n_{j'}}}/2} \]

For the Comparison of Group 3 and 4
\( \text{GH-HSD} = 4.919750 \times 0.70374 = 3.4621 \).

The 95% Confidence Interval is: \( -4 \pm 3.46222 \) \[ [-7.4621 \ -0.5379] \]

The HSD t-test is \( t_{(df=20)} = -4/0.70374 = -5.684 \), which exceeds the critical value of \( q_{[(0.95),(5.908),(4)]} = 4.919750 \). The 2-tailed p-value for the Games-Howell \( t_{(df=5.908)} = -5.684 \), is \( p = 0.0274566 \), which can be found with the SAS code: \[ \text{probq=1-PROBMC("RANGE", 5.684, . . 5.908, 4);} \]
Scheffe's Test for \( y \)

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than Tukey’s for all pairwise comparisons.

<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>2.2</td>
</tr>
<tr>
<td>Critical Value of F</td>
<td>3.09839</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>1 - 2</td>
<td>-1.0000</td>
<td>-3.7692  1.7692</td>
</tr>
<tr>
<td>1 - 3</td>
<td>0.0000</td>
<td>-2.8344  2.8344</td>
</tr>
<tr>
<td>1 - 4</td>
<td>-4.0000</td>
<td>-7.0335  -0.9665 ***</td>
</tr>
<tr>
<td>2 - 3</td>
<td>1.0000</td>
<td>-1.3404  3.3404</td>
</tr>
<tr>
<td>2 - 4</td>
<td>-3.0000</td>
<td>-5.5780  -0.4220 ***</td>
</tr>
<tr>
<td>3 - 4</td>
<td>-4.0000</td>
<td>-6.6479  -1.3521 ***</td>
</tr>
</tbody>
</table>

The Scheffé correction applied to Fisher’s LSD is simply a \( t \)-test with the variances pooled over groups as the error term. The \( df \) is simply \( df \) for the model; \( df = N-J \).

The critical value \((CV) = \sqrt{(J-1)F_{[(1-\alpha),(J-1),(N-J)]}}\). With \( \alpha = 0.05, N = 24, J = 4 \) groups,

\[
F_{[0.95,3,20]} = 3.09839 \text{ and the Scheffé Critical Value } = \sqrt{(3)3.09839} = 3.0488
\]

SAS code: \( \text{CVscheffe} = (3*(\text{Finv (.95,3,20)}))**.5; \)

For the Pairwise Comparison of Group 3 and 4, the Standard Error = \( \sqrt{\frac{2.2}{7} + \frac{1}{5}} = 0.8685 \).

The 95% Confidence Interval is: \(-4 \pm (3.0488 \times 0.8685)\) \([-6.6479, -1.3521]\]

The t-test is \( t_{(df=20)} = -4/0.8685 = -4.606 \), which exceeds the critical value of 3.0488.

The p-value can be obtained by reversing the process. If \( t_{(df=N-J)} > \sqrt{(J-1)F_{[(1-\alpha),(J-1),(N-J)]}} \) is statistically significant at \( \alpha \), then \( t_{(df=N-J)} \times (J-1) > F_{[(1-\alpha),(J-1),(N-J)]} \) is statistically significant at \( \alpha \). Thus, the p-value is \( p = 0.0020028 \), which can be obtained from the SAS code:

\[
\text{probsch}=1-\text{probf}((4.605662**2)/3,3,20);
\]

The Brown-Forsythe Adjustment, not available in SAS or SPPS, uses the unequal variance

\[
\text{SE} = \sqrt{\frac{1.1547^2}{7} + \frac{2^2}{5}} = 0.995 \text{ and the Welch-Satterthwaite } df = 5.908. \text{ The } F_{[0.95,3,5.908]} = 4.80445 \text{ and the Scheffé Critical Value } = \sqrt{(3)4.80445} = 3.7964918
\]

SAS code: \( \text{CVscheffe} = (3*(\text{Finv (.95,3,5.908)}))**.5; \)

The 95% Confidence Interval is: \(-4 \pm (3.7965 \times 0.995)\) \([-7.7777, -0.2225]\]

The t-test is \( t_{(df=5.908)} = -4/0.995 = -4.019 \), which exceeds the critical value of 3.7964918.

The p-value is \( p = 0.0396625 \), which can be obtained from the SAS code:

\[
\text{probsch}=1-\text{probf}((4.019**2)/3,3,5.908);
\]
Summary of the Pairwise Comparison of Groups 3 and 4; Mean Difference = -4.

<table>
<thead>
<tr>
<th>Method</th>
<th>Variance</th>
<th>SE</th>
<th>Critical Value</th>
<th>df</th>
<th>Half-Width</th>
<th>95 % CI Lower</th>
<th>95 % CI Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fisher’s LSD</td>
<td>Pooled</td>
<td>0.8685</td>
<td>2.086</td>
<td>20</td>
<td>1.81165</td>
<td>-5.8117</td>
<td>2.1883</td>
</tr>
<tr>
<td>Bonferroni</td>
<td>Pooled</td>
<td>0.8685</td>
<td>2.92712</td>
<td>20</td>
<td>2.5422</td>
<td>-6.5422</td>
<td>-1.4578</td>
</tr>
<tr>
<td>Šidák</td>
<td>Pooled</td>
<td>0.8685</td>
<td>2.91761</td>
<td>20</td>
<td>2.5339</td>
<td>-6.5339</td>
<td>-1.4661</td>
</tr>
<tr>
<td>Tamhane</td>
<td>Separate</td>
<td>0.9950</td>
<td>3.8705</td>
<td>5.908</td>
<td>3.8520</td>
<td>-7.8520</td>
<td>-0.1480</td>
</tr>
</tbody>
</table>

$q$-based

<table>
<thead>
<tr>
<th>Method</th>
<th>Variance</th>
<th>SE</th>
<th>Critical Value</th>
<th>df</th>
<th>Half-Width</th>
<th>95 % CI Lower</th>
<th>95 % CI Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tukey-Kramer HSD</td>
<td>Pooled</td>
<td>0.6142</td>
<td>3.95829</td>
<td>20</td>
<td>2.4309</td>
<td>-6.4309</td>
<td>-1.5691</td>
</tr>
<tr>
<td>Games-Howell</td>
<td>Separate</td>
<td>0.70324</td>
<td>4.91975</td>
<td>5.908</td>
<td>3.4621</td>
<td>-7.4621</td>
<td>-0.5379</td>
</tr>
</tbody>
</table>

$F$-based

<table>
<thead>
<tr>
<th>Method</th>
<th>Variance</th>
<th>SE</th>
<th>Critical Value</th>
<th>df</th>
<th>Half-Width</th>
<th>95 % CI Lower</th>
<th>95 % CI Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheffé</td>
<td>Pooled</td>
<td>0.8685</td>
<td>3.0488</td>
<td>20</td>
<td>2.6479</td>
<td>-6.6479</td>
<td>-1.3521</td>
</tr>
<tr>
<td>Brown-Forsythe</td>
<td>Separate</td>
<td>0.9950</td>
<td>3.7965</td>
<td>5.908</td>
<td>3.7775</td>
<td>-7.7777</td>
<td>-0.2225</td>
</tr>
</tbody>
</table>
Scheffé method applies to the set of estimates of all possible contrasts among the factor level means, not just the pairwise differences, so for a researcher only doing pairwise comparisons, the Scheffé method is an overcorrection. So what other contrasts could be of interest?

Helmert contrasts, often called “test and toss,” compare the first level of the factor with all later levels, the second level with all later levels, the third level with all later and so forth. In our \( J=4 \) group ANOVA. The null hypotheses for the Helmert contrasts are

\[
H_{0(1)}: \psi_{(1)} = [ 3 \mu_1 - 1 \mu_2 - 1 \mu_3 - 1 \mu_4 ] = 0 \\
H_{0(2)}: \psi_{(2)} = [ 0 \mu_1 + 2 \mu_2 - 1 \mu_3 - 1 \mu_4 ] = 0 \\
H_{0(3)}: \psi_{(3)} = [ 0 \mu_1 + 0 \mu_2 + 1 \mu_3 - 1 \mu_4 ] = 0
\]

Note that in a balanced design these contrasts are considered orthogonal. The dot product of each pairwise contrast is zero. For example, the dot product of contrasts 1 and 2 is 
\((3x0)+(-1x2)+(-1x-1)+(-1x-1) = 0\). Also note that this could be used as a coding scheme to represent group membership in a regression model approach to ANOVA. That is,

<table>
<thead>
<tr>
<th>Group</th>
<th>( H_1 )</th>
<th>( H_2 )</th>
<th>( H_3 )</th>
<th>Orthogonality in a statistical context implies that the Pearson correlations among ( H_1 ), ( H_2 ), and ( H_3 ) are equal to zero in a balanced design.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td></td>
</tr>
</tbody>
</table>

Polynomial contrast are used to assess trends in the data and often called trend coefficients. Suppose Groups 1 through 4 represent 0, 10, 20, 30 mg of a Drug, respectively. Then Polynomial contrast can be used to assess the form of a dosage curve.

In our \( J=4 \) group ANOVA. The null hypotheses for the Polynomial contrasts are

\[
H_{0(L)}: \psi_{(L)} = [-3 \mu_1 - 1 \mu_2 + 1 \mu_3 + 3 \mu_4 ] = 0 \text{ Linear} \\
H_{0(Q)}: \psi_{(Q)} = [ 1 \mu_1 - 1 \mu_2 - 1 \mu_3 + 1 \mu_4 ] = 0 \text{ Quadratic} \\
H_{0(C)}: \psi_{(C)} = [-1 \mu_1 + 3 \mu_2 - 3 \mu_3 + 1 \mu_4 ] = 0 \text{ Cubic}
\]

Note that in a balanced design these contrasts are considered orthogonal. The dot product of each pairwise contrast is zero. For example, the dot product of contrasts 1 and 2 is 
\((-3x1)+(-1x-1)+(1x-1)+(3x1) = 0\).

Also note that this could be used as a coding scheme to represent group membership in a regression model approach to ANOVA. That is,

<table>
<thead>
<tr>
<th>Group</th>
<th>( T_1 )</th>
<th>( T_2 )</th>
<th>( T_3 )</th>
<th>Orthogonality in a statistical context implies that the Pearson correlations among ( P_1 ), ( P_2 ), and ( P_3 ) are equal to zero in a balanced design.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>1</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
<td>-1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>-1</td>
<td>-3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
For the linear polynomial contrast the value, \( \hat{\psi}_{(L)} = [-3(2) - 1(3) + 1(2) + 3(6)] = 11 \).

The Pooled Standard Error uses the general formula:

\[
SE(\hat{\psi}) = \sqrt{\frac{\sum J c_j^2}{\sum j n_j}} \text{ for the Linear Polynomial contrast: } SE(\hat{\psi}) = \sqrt{\frac{22}{4} + \frac{1^2}{8} + \frac{1^2}{7} + \frac{3^2}{5}}
\]

\[
= 4.317857143 \text{ } SE(\hat{\psi}) = \sqrt{2.2}[4.318] = 3.082
\]

The Scheffé critical value (CV) = \( (1 - \alpha, (J-1), (N-J)) \). With \( \alpha = 0.05 \), \( N = 24 \), \( J = 4 \) groups, \( F_{[0.95,3,20]} = 3.09839 \) and the Scheffé Critical Value = \( \sqrt{(3)}3.09839 = 3.0488 \).

SAS code: CVscheffe= (3*(Finv(.95,3,20)))**.5;

The t-test is \( t_{(df=20)} = 11/3.082 = 3.57 \), which exceeds the critical value of 3.0488.

The p-value can be obtained by reversing the process. If \( t_{(df=N-J)} > \sqrt{(J-1)F_{[(1-\alpha), (J-1), (N-J)]}} \) is statistically significant at \( \alpha \), then \( t_{(df=N-J)}^2 > F_{[(1-\alpha), (J-1), (N-J)]} \) is statistically significant at \( \alpha \). Thus, the p-value is \( p = 0.0178024 \), which can be obtained from the SAS code:

probsch=1-probf(((3.57**2) / 3), 3, 20);

The Brown-Forsythe Adjustment, not available in SAS or SPPS, uses the unequal variance.

The Pooled Standard Error uses the general formula:

\[
SE(\hat{\psi}) = \sqrt{\frac{\sum J c_j^2 s_j^2}{\sum j n_j}} \text{ for the Linear Polynomial contrast: }
\]

\[
SE(\hat{\psi}) = \sqrt{\left[\frac{-3^2(2.309)^2}{4} + \frac{-1^2(0.756)^2}{8} + \frac{1^2(1.1547)^2}{7} + \frac{3^2(2)^2}{5}\right]} = \sqrt{19.45775} = 4.411
\]

and the Welch-Satterthwaite \( df = 6.2134 \) using the general formula:

\[
v = \frac{\left[\sum J c_j^2 s_j^2\right]^2}{\sum j c_j^2 s_j^2 n_j} = (19.45775)^2 / 60.933 = 6.2134
\]

. The \( F_{[0.95,3,6.2134]} = 4.654457 \) and the Scheffé Critical Value = \( \sqrt{(3)}4.6544567 = 3.73676 \).

SAS code: CVscheffe= (3*(Finv(.95,3,6.2134)))**.5;

The t-test is \( t_{(df=5.908)} = 11/4.411 = 2.494 \), which Does NOT the critical value of 3.7964918.

The p-value is \( p = 0.1983 \), which can be obtained from the SAS code:

probsch=1-probf(((2.494**2) / 3), 3, 6.2134);
Although Polynomial and Helmert-type contrasts can be performed after the rejection of the omnibus $F$-test (i.e., post-hoc), many researchers would consider these planned (or a priori) contrast. This is because if researchers were conducting a dosage-type study, they may know that are interested in linear and non-linear trends before collecting the data. This same logic may be the case for Helmert-type contrasts. The advantage of planned comparisons is that in terms of adjusting alpha the researcher is not heavily penalized for considering every pairwise comparison (i.e., Tukey-Kramer, Games-Howell) or every possible contrast (Scheffé, Brown-Forsythe). In fact many statisticians contend that if one is conducting planned contrasts, then the testing of the omnibus F-test should be bypassed! Keeping this in mind, another advantage of the planned contrast, is that it can save degrees-of-freedom and reduce the number of tests conducted, which in turn will make the correction for multiple testing less severe. For example, if researchers were to conduct a dosage study with $J=6$ doses, there would be $k=15$ pairwise comparisons. Many of the pairwise comparisons would not be of interest. **So why correct for a test that is not of interest?** In this scenario, there are $k=5$ polynomial contrast, and the researchers may only be interested in the linear (trajectory) and quadratic (leveling-off) trends, so why conduct the higher-order polynomials when they are not of interest? In this case, the researchers could bypass the omnibus F-test and analyze the $k=2$ planned contrasts. An advantage of the polynomial contrasts being orthogonal is that in a balanced design, the contrasts are uncorrelated and can be considered independent. Therefore, the Šidák correction is appropriate and does not overcorrect the adjusted alpha.

There are other types of planned contrasts. For example Dunnett-type contrasts designate a “control” or “reference” group and compare each other group to the “control” in a pairwise fashion. There are $J-1$ of there comparisons and this can be accomplished via a regression model with dummy or indicator codes.

In the $J=4$ group ANOVA, if group 1 is the “control” or “reference” group, the null hypotheses for the Dunnett contrasts are:

- $H_{0(1)}$: $\psi_{(1)} = [1 \mu_1 - 1 \mu_2 + 0 \mu_3 + 0 \mu_4] = 0$
- $H_{0(2)}$: $\psi_{(2)} = [1 \mu_1 + 0 \mu_2 - 1 \mu_3 + 0 \mu_4] = 0$
- $H_{0(3)}$: $\psi_{(3)} = [1 \mu_1 + 0 \mu_2 + 0 \mu_3 - 1 \mu_4] = 0$

Note that these contrasts are NOT orthogonal. The dot product of each pairwise contrast is NOT zero. For example, the dot product of contrasts 1 and 2 is $(1x1)+(-1x0)+(0x-1)+(0x0) = 1$. Also note that this could be used as a coding scheme to represent group membership in a regression model approach to ANOVA. Dummy coding can also be used to accomplish Dunnett-type Tests.

<table>
<thead>
<tr>
<th>Group</th>
<th>Effect Coding</th>
<th>Dummy Coding</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$E_1$</td>
<td>$E_2$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
data bal;set bal;
if group = 1 then do;
D2 = 0;D3 = 0;D4 = 0;end;
if group = 2 then do;
D2 = 1;D3 = 0;D4 = 0;end;
if group = 3 then do;
D2 = 0;D3 = 1;D4 = 0;end;
if group = 4 then do;
D2 = 0;D3 = 0;D4 = 1;end;run;

proc reg data=bal;
model y = D2 D3 D4;run;

proc glm data=bal;class group;
model y = group;lsmeans group
/adjust=dunnett tdiff pdiff;
contrast "Dunnet12" group 1 -1 0 0;
contrast "Dunnet13" group 1 0 -1 0;
contrast "Dunnet14" group 1 0 0 -1;run;

The REG Procedure
Dependent Variable: y
Analysis of Variance

                      Source      DF    Squares    Square   F Value  Pr > F
Model                   3  55.333333  18.444444    8.38   0.0008
Error                   20  44.000000   2.200000
Corrected Total        23  99.333333

Root MSE     1.48324    R-Square    0.557047
Dependent Mean   3.16667    Adj R-Sq    0.4906
Coeff Var      46.83915

Parameter Estimates

                            Parameter     Standard

                                Variable      DF      Estimate      Error    t Value  Pr > |t|
Intercept                 1       2.000000   0.741620    2.70      0.0139
D2                       1       1.000000   0.908300    1.10      0.2840
D3                       1   9.37618E-16    0.929670    0.00      1.0000
D4                       1       4.000000   0.994990    4.02      0.0007

The GLM Procedure
Dependent Variable: y

                          Source      DF    Squares    Mean Square   F Value  Pr > F
Model                   3  55.333333  18.44444444    8.38   0.0008
Error                   20  44.000000  2.20000000
Corrected Total        23  99.333333

R-Square    Coeff Var   Root MSE   y Mean
0.557047   46.83915    1.483240   3.166667

Least Squares Means

Adjustment for Multiple Comparisons: Dunnett
H0:LSMean=Control

group     y LSMEAN   t Value  Pr > |t|  
1        2.00000000    2.00000000   1.10      0.5334  t-tests are the same
2        3.00000000    3.00000000   2.00000000  0.00      1.0000  p-values have been
3        6.00000000    6.00000000   4.02      0.0018  adjusted
4

Contrast     DF    Contrast SS    Mean Square    F Value   Pr > F
Dunnet12     1  2.66666667  2.66666667    1.21   0.2840  
Dunnet13     1   0.00000000  0.00000000   0.00   1.0000  
Dunnet14     1 35.55555556 35.55555556    16.16   0.0007