1 Multiple Comparisons: examples

We will use the cuckoo egg data set to exemplify various multiple comparison procedures. Examples will include 1) Two *a priori* orthogonal contrasts, 2) Tukey's and Fisher's methods for pairwise differences, and 3) Scheffe's method for a *post hoc* contrast.

1.1 ANOVA results

Since these tests will rely on results from the ANOVA of the cuckoo egg data, here are the results of that analysis:

```
                              The GLM Procedure
Dependent Variable: length
                                     Sum of
 Source                      DF     Squares    Mean Square   F Value   Pr > F
 Model                        5  43.2928889      8.6585778      9.58   <.0001
 Error                       84  75.9200000      0.9038095
 Corrected Total             89 119.2128889

              R-Square     Coeff Var      Root MSE    length Mean
              0.363156      4.221740      0.950689       22.51889

 Source                      DF   Type III SS    Mean Square   F Value   Pr > F
 species                      5   43.29288889     8.65857778      9.58   <.0001

                              The GLM Procedure
               Level of           ------------length-----------
               species       N            Mean          Std Dev
               HSprw        15      23.1700000       1.04690019
               MPipit       15      22.2766667       1.15601944
               PWtail       15      22.9033333       1.06761862
               Robin        15      22.5433333       0.69638522
               TPipit       15      23.0900000       0.90142744
               Wren         15      21.1300000       0.74373574
```

1.2 Two *a priori* orthogonal contrasts

Suppose before data collection has occurred, it was desired to test i) whether the mean cuckoo egg length differed between Tree Pipit nests and Meadow Pipit nests, and also ii) whether mean cuckoo egg length for Pipit nests (Tree and Meadow averaged) differed from Robin nests. Then the two contrasts are:

$$L_1 = \mu_{TP} - \mu_{MP} \text{ and } L_2 = (\mu_{TP} + \mu_{MP})/2 - \mu_R.$$

Here μ_{TP}, for example, is the population mean of the Tree Pipit (TP) group. Since they were specified *a priori* and they are orthogonal (check this as an exercise), we can use separate t tests for each contrast. We use $\alpha = .05$ separately for each test. We get

$\widehat{L}_1 = \overline{y}_{TP} - \overline{y}_{MP} = 23.09 - 22.27 = .82$, and $\widehat{L}_2 = (\overline{y}_{TP} + \overline{y}_{MP})/2 - \overline{y}_R = (23.09 + 22.27)/2 - 22.54 = .14$.

and also

$$\widehat{Var}(\widehat{L}_1) = \text{ MSE } \sum_{i=1}^{t} \frac{a_{1i}^2}{n_i} = (.904)((1)^2/15 + (-1)^2/15) = .121, \text{ and}$$

$$\widehat{Var}(\widehat{L}_2) = \text{ MSE } \sum_{i=1}^{t} \frac{a_{2i}^2}{n_i} = (.904)((\frac{1}{2})^2/15 + (\frac{1}{2})^2/15 + (-1)^2/15) = .0904.$$

Then for $H_0 : L_1 = 0$ we have t $= \widehat{L}$ / *s.e.*($\widehat{L}$) $= .82/\sqrt{.121} = 2.36$, and for $H_0 : L_2 = 0$ we have t $= \widehat{L}$ / *s.e.*($\widehat{L}$) $= .14/\sqrt{.0904} = .47$. The degrees of freedom for MSE is 84 so if we are conservative and use $df = 80$ we get $t_{80,.975} = 1.99$. Thus we reject $H_0 : L_1 = 0$ in favor of $H_A : L_1 \neq 0$, but fail to reject $H_0 : L_2 = 0$. Equivalently, we can use the CONTRAST or ESTIMATE statement in PROC GLM to get a P value of .02 for $H_0 : L_1 = 0$ and a P value of .64 for $H_0 : L_2 = 0$.

```
                              The GLM Procedure
Dependent Variable: length
Contrast                     DF     Contrast SS     Mean Square    F Value    Pr > F
Tree vs Meadow Pipits         1      4.96133333      4.96133333       5.49    0.0215
Pipits vs Robins              1      0.19600000      0.19600000       0.22    0.6426

                                                   Standard
  Parameter                    Estimate               Error     t Value      Pr > \vert t\vert
  Tree vs Meadow Pipits      0.81333333          0.34714253        2.34        0.0215
  Pipits vs Robins           0.14000000          0.30063425        0.47        0.6426
```

1.3 Tukey's and Fisher's methods

For Tukey's method we use

$$W = q_{t,df,1-\alpha} \sqrt{\frac{\text{MSE}}{n}}, \text{while for Fisher's LSD method we use } LSD = t_{df,1-\alpha/2} \sqrt{\text{MSE}(\frac{1}{n_i} + \frac{1}{n_{i'}})}.$$

The Studentized range term for W is now approximated using $df = 60$ to get $q_{6,60,.95} = 4.16$, and the t value for LSD is still $t_{80,.975} = 1.99$. Thus we have

$$W = 4.16\sqrt{\frac{.904}{15}} = 1.02 \text{ and } LSD = 1.99\sqrt{.904(\frac{1}{15} + \frac{1}{15})} = .69.$$

Any two sample means that differ by 1.02 for Tukey or .69 for Fisher's LSD lead us to declare their population means different. See the SAS output for an illustration of the results.

The GLM Procedure

t Tests (LSD) for length

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

Alpha	0.05
Error Degrees of Freedom	84
Error Mean Square	0.90381
Critical Value of t	1.98861
Least Significant Difference	0.6903

Means with the same letter are not significantly different.

t Grouping		Mean	N	species
	A	23.1700	15	HSprw
	A			
	A	23.0900	15	TPipit
	A			
B	A	22.9033	15	PWtail
B	A			
B	A	22.5433	15	Robin
B				
B		22.2767	15	MPipit
	C	21.1300	15	Wren

The GLM Procedure

Tukey's Studentized Range (HSD) Test for length

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	84
Error Mean Square	0.90381
Critical Value of Studentized Range	4.12462
Minimum Significant Difference	1.0125

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	species
A	23.1700	15	HSprw
A			
A	23.0900	15	TPipit
A			
A	22.9033	15	PWtail
A			
A	22.5433	15	Robin
A			
A	22.2767	15	MPipit

```
                 B        21.1300      15     Wren
```

Note in the SAS output that the values for W and LSD are slightly different than ours, because they do not have to approximate the degrees of freedom from the MSE.

1.4 Scheffe's method

Suppose that after seeing the data, we decide to compare the mean length of cuckoo eggs in Wren nests to all other species combined. Then $L = \mu_W - (\mu_{TP} + \mu_{MP} + \mu_R + \mu_P + \mu_H)/5$, where W = Wrens, TP = Tree Pipits, MP = Meadow Pipits, R = Robins, P = Pied Wagtails, and H = Hedge Sparrows. To test $H_0 : L = 0$ as a *post hoc* contrast, we use Scheffe's method. Then we have

$\widehat{L} = \overline{y}_W - (\overline{y}_{TP} + \overline{y}_{MP} + \overline{y}_R + \overline{y}_P + \overline{y}_H)/5 = 21.13 - (23.09 + 22.27 + 22.54 + 22.90 + 23.17)/5 = 1.66$ and

$$\widehat{Var}(\widehat{L}) = \text{ MSE } \sum_{i=1}^{t} \frac{a_i^2}{n_i} =$$

$$(.904)\ ((1)^2/15 + (-\frac{1}{5})^2/15 + (-\frac{1}{5})^2/15 + (-\frac{1}{5})^2/15 + (-\frac{1}{5})^2/15 + (-\frac{1}{5})^2/15) = .072$$

We need to compare $\widehat{L}$ to Scheffe's $S = \sqrt{(t-1)F_{t-1,df,1-\alpha}}\sqrt{\widehat{Var}(\widehat{L})}$. Here we approximate the 84 degrees of freedom with $df = 80$, and get

$$S = \sqrt{(t-1)F_{t-1,df,1-\alpha}}\sqrt{\widehat{Var}(\widehat{L})} = \sqrt{5F_{5,80,.95}}\sqrt{.072} = \sqrt{5(2.33)}\sqrt{.072} = (3.41)(.27) = .92.$$

Since $|\widehat{L}| > S$, we reject $H_0 : L = 0$, and conclude that $\mu_W - (\mu_{TP} + \mu_{MP} + \mu_R + \mu_P + \mu_H)/5 \neq 0$. Unfortunately, SAS only implements Scheffe's method for pairwise contrasts, for which it is too conservative to use.