1 Visualization, checking assumptions for ANOVA

As with regression analysis, it is important both to visualize the data and to assess the assumptions of the ANOVA model. We can use boxplots to visualize data for ANOVA models, and we can look at residual-by-predicted plots and normal plots to check the assumptions of homogeneous variance and normality. See the computer programs for the cuckoo data for details.

2 Multiple Comparisons

If the ANOVA null hypothesis, $H_0 : \mu_1 = \mu_2 = ... = \mu_k$ is rejected, we often wish to learn more about which populations means differ. For example, for the cuckoo egg data, we may wish to further address which species have different mean cuckoo egg lengths in their nests. This involves testing other hypotheses (such as $H_0 : \mu_1 = \mu_2$) that involve a subset of the group means. The concern then arises about controlling the Type I error rate if we make many such tests. There is a vast literature concerning multiple comparison tests, enough to generate entire texts and to occupy semester-length courses. We will cover only a small amount of this material, consisting of a few common procedures. Three important questions that we must address when performing multiple comparison tests are:

1. Did we propose the hypotheses to be tested before collecting the data (*a priori*), or after looking at the data (*post hoc*)?

2. If the comparisons are *a priori*, do they involve independent (orthogonal) comparisons?

3. Are the comparisons pairwise, or do they involve more complicated hyptheses (such as $H_0 : (\mu_1 + \mu_2)/2 = \mu_3$)?