Ch 23 – Two-factor studies—unequal sample sizes

Consequences of unequal sample sizes:

- In the ANOVA, the sums of squares depend on order of entry of terms into the model, so base tests on general linear test approach.
- Main effect treatment contrasts estimates are more complicated, whether the model includes interactions or not

Example: p954, ch23eg.txt

Consider the analysis with the standard two-factor model with interactions:

Discuss ANOVA and the general linear test approach, using Type III tests initially at least.

Recall for the LSEs: $\hat{\mu}_{ij} = \bar{y}_{ij.}$. So, for a main-effect-of-A treatment contrast for example, the LSE of $\sum_i c_i \bar{\mu}_{i.} = \sum_i c_i (\sum_j \bar{y}_{ij.}/b) \not\equiv \sum_i c_i \bar{y}_{i..}$, i.e. these are not equivalent for unequal sample sizes. For PROC GLM in SAS, the LSMEANS statement and the ESTIMATE statement will use the LSEs, as should be done, whereas the MEANS statement will be okay for cell means but not for main effect contrasts.

Tukey's method is still applicable for comparing all ab treatment effects, and also for comparing all levels of either factor pairwise. The methods of Scheffé and Bonferroni methods are also applicable.

What if the main effects model is used?

LSEs are not easily expressed algebraically for unequal sample sizes, so rely on software to do the computations.

For ANOVA, sums of squares depend on the order in which terms are entered into the model.

Tukey's method is probably okay to use in this context, but it remains an open problem to show it is exact or conservative in this context, since factor level effects are generally correlated.

Checking model assumptions: No change in use of residual plots.

Homework: For the data of problem 23.9, analyze the data using the full model. In particular: (a) use residuals to evaluate model assumptions; (b) plot yhat*A=B and yhat*B=A and discuss what you anticipate statistical analysis may show; (c) use ANOVA to test for interactions and for each main effect, adjusting appropriately in each case; (d) Use the Bonferroni method of multiple comparisons for each main effect so as to have at least 90% confidence in all pairwise comparisons, and discuss the results.