Lab 10 - Two Way ANOVA

April 17, 2017

In this lab we go over Two-Way Balanced and Unbalanced ANOVA in R.

The data come from a study on hospitalization of patients with kidney failure. These patients are commonly treated on dialysis machines that filter toxic substances from the blood. The appropriate "dose"; for effective treatment depends, among other things, on duration of dialysis treatment and weight gain between treatments as a result of fluid buildup. Duration was categorized into two groups, short (less than four hours) and long (four or more hours) coded as 1 and 2 respectively. Average weight gain was categorized into slight, moderate, and severe coded as 1, 2, and 3 respectively. The response variable of interest is the number of days hospitalized due to subsequent kidney disease. Prior analysis indicated that transforming the number of hospital days to LOG=LOG(DAYS+1) made the data more normal, and equalized the variance.

To study the effects of duration of dialysis treatment and weight gain on the number of days hospitalized due to disease, a random sample of 10 patients per group who had undergone treatment at a large dialysis facility was obtained. This full data set has equal numbers of observations in each cell, so the design is balanced. The unbalanced design is obtained by dropping the last two patients in the long duration and severe weight gain category, and is done purely for illustrative reasons. In a real study, experimental units might be missing (due to causes unrelated to treatment, such as moving out of the area), or might be dropped because they are outliers. In an observational study such as this one, unequal sample sizes may also occur because not enough experimental units are available in some treatment groups.

a) Fit an interaction Two-Way ANOVA model. Check to see that for both the balanced and unbalanced designs, the fitted values are just the cell means.

b) Now fit just the main effects model. Notice that for the balanced design, the type I and type III sums of squares are the same, but for the unbalanced design, they differ. Why is this?

c) In the balanced data, which effects are significant? What is the appropriate reduced model for this data? Do this by first testing for an interaction, and then testing for individual effects if no interaction is present. Since any of these results would be scientifically significant, be sure to test for multiple tests.

d) Regardless of the model chosen in part C, continue with the interaction model. Estimate main effects and interactions for the balanced model for each factor level and each treatment. How would you get confidence intervals or perform tests for the indvidual main effects and interaction effects?

e) Construct the ANOVA table for the reduced model.

f) Contrasts can be completed as they are done in One-Way Anova. Create a contrast for testing Slight WeightGain against Moderate to Severe WeightGain, and Moderate WeightGain against Severe WeightGain.