Stat 200B Spring 2000

Lab 5

Due Friday 21 Apr

In this lab we shall explore some fractional factorial experiments. Consider first the data file `l5.dat`. This data file consists of five columns the first four columns are the treatments (A,B,C,D) at which an experiment has taken place, coded so that "-1" is equivalent to the low level of that factor and "+1" is the high level of that factor. The final column is the response variable in the experiment. You will note that there are only 8 replications of the experiement, thus we have a half fraction of the entire factorial experiment. Why would one choose to carry out a fractional factorial experiment rather than the full factorial experiment? For this data set can you determine what fraction has been taken? (it may not be immediately obvious and you may need to complete the rest of the lab before you can answer this question). In any case read in the data and carry out an ANOVA on the data set. Consider relating the response to the four explanatory variables (in a linear fashion). Produce an ANOVA table and interpret what you see. Which variable seems to be most explanatory? Consider ANOVA including higher order effects. What do you notice when you do this? Explain if possible.

Now suppose that in fact we did in fact have results of the experiement at all the different treatment level combinations (as luck would have it we do). The file `l5b.dat` contains observations at all 16 treatment-level combinations. Perform the same ANOVA you performed on the fractional data. What do you observe? Is this different from what you expected? Consider adding in higher order terms into your ANOVA. What do you see now? Do any of your results now reconcile with what you saw earlier? Explain if possible. Do you think that the original fraction was a sensible fraction to take? Suggest a better fraction to take and give a reason why. Consider subsetting your data with this fraction and carrying out the ANOVA in this fraction. What do you see now?

Make sure to clearly state your assumption through out the report.

You may find `aov()`, `summary()`, `summary.lm()` useful. There is material in chapter 6 of Venables and Ripley that may also be appropriate.

NOTE: the data files are located in `/saruman/class/data/s200b/data/2000`