

Stat 415 - Mixed Linear Models, **Homework**

Due date: Friday, Apr 22, 5 pm, in my office: 120 Snedecor

Guidelines:

You are encouraged to work together. However, please write your answers individually.

There are two problems. Both describe a study, then ask you to analyze a data set and make conclusions. The data files for that problem are on the stat415 part of my web site (<http://www.public.iastate.edu/~pdixon/stat415>). Files (data file, SAS code, and R code) for the class examples (spruce growth, fungicide resistance, and parachute line strength) are also on my web site.

Problem 1. The following data come from the National Youth Survey. One of the questions asked in that survey provides information about tolerance of deviant behavior. Deviant behavior, for this survey, includes things like cheating, destroying the property of others, using alcohol, and other characteristics. These are combined into an overall score for 'tolerance of deviant behavior'. Higher values indicate more tolerance. There are five measurements of tolerance for each of the 16 people in the data set.

The data are in tolerance.txt. The columns are:

id: identifies each person. Different value for each person

male: 0 for female, 1 for male

tolerance: the measured tolerance score. There are five observations for each person.

The goals of the analysis are to:

estimate the variability between people and variability between measurements

estimate the difference in mean tolerance between males and females.

- a) Write down a model for the observations. Your model should be appropriate for the goals of the study. Define the terms and parameters in your model.
- b) What parameter (or parameters) in your model describes the:
 - variability between people
 - variability between measurements
 - the difference in mean tolerance between males and females.
- c) Estimate the variance components
- d) Estimate the difference between males and females, estimate the s.e. of that difference, and test whether the true difference is 0.

Problem 2. What parts of the brain are responsible for memory? Various studies on folks with amnesia suggest that short-term memory (i.e. cramming for an exam) is stored in a different part of the brain from long-term memory (i.e. the good stuff). Studies of brain-injured patients suggest that folks with a damaged hippocampal formation lose short-term memory but retain long-term memory. These data come from an experimental study on monkeys to evaluate this.

18 monkeys were taught to discriminate between pairs of objects. They are supposed to remember which is the 'good' object in each pair. Each monkey was taught 100 pairs of objects: 20 pairs at each of 16, 12, 8, 4, and 2 weeks prior to the treatment. On treatment day, the hippocampal formation was blocked ($\text{trt}=2$) in 11 randomly chosen monkeys. The remaining 7 monkeys were untreated controls ($\text{trt}=1$). Each monkey was then tested to see whether they could distinguish each pair of objects. The response is the percent correct (out of 20) for each monkey for each week of training. Because R gives correct results only for balanced data, we are only using the data from weeks 2, 4, 12, and 16. Weeks 2 and 4 are considered short term memory; weeks 12 and 16 are long-term memory.

Another description of the experiment is:

This study has a 2x2 factorial treatment structure. The investigators measured the percent correct response for all combinations of TREATMENT (2 levels, blocked or control), and MEMORY (short term or long term). Monkeys were randomly assigned to one of the two levels of TREATMENT. Weeks are replicated observations of the two levels of MEMORY. If it helps to think of weeks as "randomly assigned" to short-term or long-term memory, you may do that. There are a total of 18 monkeys and 72 observations.

The data are in learn2.txt. The columns are:

monkey: id number for each monkey

trt: TREATMENT: 1 for control, 2 for blocked hippocampal formation

memory: short or long term MEMORY

week: weeks before treatment

correct: the response, percent correct

- a) Outline the analysis of variance table, indicating sources of variation. If you can calculate the degrees of freedom for each source of variation, that is worth extra credit.
- b) There are two error terms here, one for monkey-monkey variation, and one for observation-observation variation. Indicate which is the appropriate one for testing each effect in the ANOVA table.
- c) Analyze the data. The investigators are most interested in the interaction between MEMORY (long or short) and TREATMENT. Is there evidence of an interaction?