Intermediate Applied Statistics
STAT 460

Lecture 18, 11/10/2004

Instructor:
Aleksandra (Seša) Slavković
sesa@stat.psu.edu

TA:
Wang Yu
wangyu@stat.psu.edu

Revised schedule

Nov 8 10 a.m. 2-way ANOVA
Nov 10 lecture on two-way ANOVA and blocking
Nov 12 lecture on repeated measures and review
Nov 13 lab on repeated measures
Nov 11 lecture on categorical data/logistic regression
HW due
Nov 11 lecture on categorical data/logistic regression
HW due
Nov 12 lab on logistic regression & project II introduction
Nov 10 lecture
HW 10 due
Post HW 11
Nov 14 lab
Nov 12 lecture on categorical data/logistic regression
Nov 11 lecture on categorical data/logistic regression
HW due
Nov 11 lecture on categorical data/logistic regression
HW due
Dec 3 lecture and Quiz
Nov 19 lecture on categorical data/logistic regression
Nov 17 lecture on categorical data/logistic regression
HW due
Post HW 9
Nov 15 lab on repeated measures
Nov 13 lecture repeated measure and review
Nov 10 lecture on two-way ANOVA and blocking
Post HW 8
Dec 6 lab
Dec 3 lecture
Dec 1 lecture
HW 11 due
HW 11 due
Dec 13 Project II due

Last lecture

☐ Review Two-Way ANOVA in the context of experimental design
☐ Randomized design
☐ Blocking

This lecture

☐ Quiz grades
☐ Repeated measure

Quiz 3 scores

☐ Summary:
Min. 1st Qu. Median Mean 3rd Qu. Max.
37.00 70.00 83.00 77.55 88.00 93.00

☐ The decimal point is 1 digit(s) to the right of the |
3 | 7
4 | 3
5 | 68
6 | 479
7 | 02357
8 | 2356678899
9 | 01223

Review: Blocking

☐ A blocking factor classifies subjects into groups, often according to natural differences or other relationships not directly related to the treatment of interest.

☐ A blocking factor is usually a random-effects but can sometimes be treated as fixed-effects.

☐ Blocking factors are included in the calculations to get a more precise and realistic model, but their effects are not usually of much interest in themselves.
Examples:

- Three different varieties of wheat (represented as red, blue, and gray) are compared. Each variety is planted in each of three different fields. A field is a block. There is one fixed treatment factor (variety) and one blocking factor.

Examples (contd.):

- Compare the GPA’s of the first and second-born children in many families. Family is a block, birth order is the “treatment” (not really an experiment though). This can be done with a paired t-test as well as an ANOVA.
- Study the effects of various factors on paper airplane flight. Let more than one person throw the planes. Use thrower as a blocking factor.

Review: Blocking

- Blocking is important but makes the design and analysis of an experiment more complicated.
- In the simplest case, each level of the blocking factor has at least one of its experimental units assigned to each level of the treatment factors. However, there are more complicated situations.

Special ANOVA Designs Related to Blocking

- Special Cases of Blocking
  - Blocking as Repeated Measures
  - Treatments applied directly to blocks (Split plot)
  - More than one blocking factor (e.g. Latin Square approach)
  - Each block doesn’t get all treatments (Balanced Incomplete Block designs)

Repeated Measures

- Chapter 16, Sleuth
- More than one response per subject (observational unit) is recorded
  - As a series over time
  - Over direction in space
  - Or with different treatments
- These can sometimes be treated as ANOVAs where the subject is a block.
- Examples: comparing glucose levels over time after eating different kinds of food

Repeated Measures

- Synonym: Within-subjects design
- Between-subjects design
  - Each subject receives a different treatment
  - Only one outcome measurement is made
- What assumption is violated?
  - Lack of independence of errors for pairs of measurements made on same subject
What Repeated Measures Isn’t

- If the different responses are different concepts or have different distributions
- Need MANOVA, or at least separate ANOVA’s.
- For example, if we wanted to compare the tissue mercury, cadmium, boron, and copper levels of turtles at different sites, then we probably shouldn’t treat “type of metal” as a factor. (Site is a factor and turtle is nested within site.)

Paired t-test

- The paired t-test with before and after can be seen as the simplest repeated measures ANOVA.

Situations with Repeated Measures over Time

- Longitudinal observational studies
- Crossover experiments
  - Each subject receives more than one treatment level
- Split-plot in time (for two treatment factors)

Split Plots

- One well-understood type of repeated measures design is the “split-plot” design.
- This is a design in which there are two treatments of interest. One is applied directly to blocks and the other to individual units within blocks.

A Crossover Experiment in Humans

The output below is based on a study done on predicting heart rate while doing stepping exercises. “Subject” refers to the person doing the exercises; each person was considered a separate level because there was more than one measurement per person (actually, I’m oversimplifying a little – see the web page below for more information). “Height” refers to the height that the subject had to step up to. “Speed” refers to the frequency with which the person had to step.

The output below is based on a study done on predicting heart rate while doing stepping exercises. “Subject” refers to the person doing the exercises; each person was considered a separate level because there was more than one measurement per person (actually, I’m oversimplifying a little – see the web page below for more information). “Height” refers to the height that the subject had to step up to. “Speed” refers to the frequency with which the person had to step.

This is repeated measures data (the subjects are blocks) but that is not a problem here since the measurements are considered conditionally independent rather than serially correlated.

Data are from Data and Story Library, Stepping dataset, http://lib.stat.cmu.edu/DASL/Datasets/Stepping.html. Also see http://lib.stat.cmu.edu/DASL/Stories/SteppingandHeartRates.html.

Repeated Measures

- Advantages
  - More information per observational unit
  - More power (through canceling subject variability)
  - Ability to study time trends
  - Reduced number of subjects needed

- Disadvantages
  - More complicated analysis needed (because of independence assumption)
  - Includes possible confounding
Approaches to Analyzing Repeated Measures

- And Other Multivariate Data
  - Convert to a univariate summary measure
  - Treat each time point separately (e.g. time series analysis)
  - Use a multivariate test (e.g. ANOVA)
  - Use subject as block, time as a factor

Paired t-test

- A researcher interested in Friday the 13th looked at traffic in several different months on the 6th vs. the 13th (http://lib.stat.cmu.edu/DASL/Stories/Fridaythe13th.html). The outcome is the traffic volume in a certain area between certain times. The main explanatory variable is day 6 vs. day 13 of the month.

- Handout lec18example.doc

Next

- Lab: repeated measures and review
- Categorical data
- Logistic regression