Practical Statistical Methods for Linguistics

Michaelmas Term 2010-11, weeks 3, 5, 7, Wednesday, 3-4, Centre for Linguistics (room 205)
http://kochanski.org/gpk/teaching/1011statistics   (This is http://kochanski.org/gpk/teaching/1011statistics/description.pdf)

Dr Greg Kochanski

This course exists to help you do statistical analysis on your data. Empirical methods are becoming an incresingly important part of linguistic research, and empirical methods need statistical analyses. This course should help you avoid some of the traps and introduces the R statistical software package.

One aim of the course is to prepare MPhil students who are required to undertake linguistic research for their dissertation. However, empirical methods are an increasingly important part of linguistics and all advanced students need to know something about the range of research methods which are used in their field. A second aim of this course, therefore, is to help students understand and interpret experimental and statistical results. The course is part of the core teaching for first year Linguistics MPhil students; philologists and MSt students may attend if they choose. DPhil students are also welcome to attend.

There will be no separate and specific assessment of your work for the course, but at the end of the first year MPhil students will be expected to produce a dissertation proposal which, in addition to setting out the proposed research topic and question, describes the methods you intend to employ and explains the choices you have made. This course will help you prepare that proposal. Note, however, that the proposal will need to be discussed with your dissertation supervisor before submission.

If you are planning to use quantitative methods in your research I advise that you attend the lectures provided by the Social Sciences Division. Since these must be paid for by the Faculty, you should inform Kate Dobson of your intention to attend.

PROGRAMME

1. Designing Linguistic Experiments
   - How big should my experiment be?
     - Counting statistics and the square-root-of N.
     - Paired sample tests
   - Can you disprove it?

- The Trap of the Broad Hypothesis
- Pilot Experiments
- Brief Intro to R

2. A class Example: Counting Cars
- Getting data into R
- Basic use of R
- logistic regression

3. Working examples and discussion
    - Bring in your own examples to discuss, set up or work
    - Modern techniques: Bootstrap resampling and replication

Some general/introductory reading on quantitative research methods

Gonick, L. & W. Smith (1993). *The Cartoon Guide to Statistics*. New York: Harper Collins.

Johnson, K. (2008). *Quantitative Methods in Linguistics*. Oxford: Blackwell.

Milroy, L. & M. Gordon (2003). *Sociolinguistics: Method and Interpretation* Oxford: Blackwell.

Karl R. Popper, *Conjectures and Refutations: The Growth of Scientific Knowledge*, Routledge, 2003 ISBN 0415285941

Woods, A., Fletcher, P. & A. Hughes (1986). *Statistics in Language Studies.* Cambridge: CUP.

Wray, A., K. Trott & A. Bloomer. (1988). *Projects in Linguistics*. London: Arnold. [Basic, but useful.]

Web reading for the course

http://www.kochanski.org/gpk/teaching/0910ACTL_UCL/getting_good_data.pdf

http://www.kochanski.org/gpk/teaching/0401Oxford/doubt.pdf

http://www.kochanski.org/gpk/teaching/0401Oxford/sampling.pdf

http://www.kochanski.org/gpk/teaching/1011statistics/StatisticalModels.pdf

http://www.r-project.org/ Download the R software package.

http://gking.harvard.edu/zelig/ A library to help you do standard analyses.

http://www.rseek.org/ A search engine for information on R and statistics.

http://www.stat.wisc.edu/~deepayan/SIBS2005/slides/introduction.pdf

http://www.kochanski.org/gpk/teaching/0601Oxford/Rule_of_Thumb.pdf

http://www.kochanski.org/gpk/teaching/0601Oxford/sumstats.pdf

http://www.kochanski.org/gpk/teaching/0601/02Rdemo.pdf

https://netfiles.uiuc.edu/cls/www/course/ling542/ling542.html