Introduction to Statistical Methods for Linguistics

Michaelmas Term 2009-10, weeks 2, 4, 6, 8
Thursday, 2--3.30, Centre for Linguistics (room CLP)

Dr Ros Temple & Dr Greg Kochanski

This course considers aspects of statistical design and methodology which are an important part of quantitative linguistic research, and introduces some of the basic tools which are commonly used by researchers in linguistics. One aim of the course is to prepare MPhil students who are required to undertake linguistic research for their dissertation, by enabling them to make appropriate methodological choices when planning their own projects. 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 we strongly advise that you attend the lectures provided by Dr James Tilley for 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. *Introduction to quantitative methods in linguistic research*.

This session will introduce you to the issues which need to be addressed when doing quantitative research in linguistics (or indeed any other discipline). For example, how to formulate research questions and hypotheses. Examples will be taken from variationist linguistics, but the generic issues are relevant to all quantitative approaches to linguistic research.

Sessions 2 – 4 introduce you to more technical details about statistical analysis. These sessions will involve practical work by students.

2. *Counting Statistics and Sampling*.
    - When are two counts significantly different?
    - When counts get sparse: how to deal with just a few examples.
    - Dealing with data in SPSS.
    - ANOVA

- logistic regression
- Brief intro to R.

3. *Sqrt(N) is your friend: planning the size of an experiment.*
    - How many subjects?
    - How much can you say if it's significant? If it's not?
    - Bonferroni corrections and doing more than one test.

    Choosing your statistical test:
    - Gaussian or non-Gaussian
    - Continuous data, ordered data, vs. separate categories.
    - Paired sample vs. not.
    - t-tests
    - Non-parametric relatives of t-tests

4. *Linear regression and the stuff you do to prepare for it.*
    - Students bring in problems / request reprise of topics

    If your data is too rich:
    - Principle component analysis
    - Multidimensional Scaling

    Modern statistics that you should be aware of:
    - Monte-Carlo simulation
    - Bootstrap Resampling

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.

Tagliamonte, S. (2006). *Analyzing Sociolinguistic Variation.* Cambridge: CUP.

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. [VERY basic, but contains a lot you need to know, so useful if you're starting from scratch]

http://www.uvm.edu/ets/statistics/SPSS15Manuals/SPSS Brief Guide 15.0.pdf

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/doubt.pdf

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

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

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