Statistics 149 – Introduction to Probability – Spring 2006

Instructor: Mark Irwin
Office: Science Center 611
Phone: 617-495-5617
e-mail: irwin@stat.harvard.edu
Web site: http://www.courses.fas.harvard.edu/~stat149/

Office Hours: Monday 1:00 – 2:00,
Thursday 2:00 – 3:00,
or by appointment

Teaching Fellow: Alan Lenarcic
e-mail: lenarcic@fas.harvard.edu

Office Hour: To be announced

Objectives

An introduction to methods for analyzing categorical data. Emphasis will be on understanding models and applying them to datasets. Topics include visualizing categorical data, analysis of contingency tables, odds ratios, log-linear models, generalized linear models, logistic regression, Poisson regression and model diagnostics. Examples drawn from many fields, including biology, medicine and the social sciences.

Prerequisites

Statistics 139 or equivalent.

Lectures

Tuesday & Thursday, 10:00 – 11:30, Science Center 111

Sections

To be scheduled

Required Text

Ramsey FL and Schafer DW (2002). The Statistical Sleuth: A Course in Methods of Data Analysis, 2nd edition. Duxbury.

Other References

McCullagh P and Nelder JA (1989) Generalized Linear Models, 2nd edition. Chapman and Hall.

Agresti A (1996). An Introduction to Categorical Data Analysis. Wiley.

Hosmer D and Lemeshow S (2000). Applied Logistic Regression 2nd edition. Wiley.

Dobson AJ (1990). An Introduction to Generalized Linear Models. Chapman & Hall

Computing:

The statistical software package that will be supported in this course is R. Please refer to <http://www.r-project.org/> for instructions on how to download R. Discussion sections will cover basic instructions on R. There is also information about R available at <http://www.people.fas.harvard.edu/~mirwin/computing/splus.html>, which can be accessed via the course web site. You are free to use whatever other software that you want as long as it can handle the problems in this course. If you use another software package, the TF may not by able to assist you with your computing program.

Grading:

Homework (30%)

Homework assignments will be posted on the course page, in addition to being distributed in lecture. Please show work supporting your answers. Correct answers without supporting work will not receive credit. There will be 5 or 6 assignments in total. Late homework will not be accepted.

Final Project (10%)

A data analysis project.

Midterm (25%)

Tuesday, March 21st, in class (Tentative).

Final Exam (35%)

Exam Group: 12, 13. Date and time to be announced.