

Sociology V3212: Statistics/Methods

December 20, 2005

Tuesday and Thursday 10:35am-11:50am
644 Seeley Mudd Bldg.

	Instructor	TA
	Aaron Gullickson	Ho-Dae Chong
office:	412 Fayerweather Hall	North Foyer of Fayerweather, (4th floor)
email:	ag2319@columbia.edu	hc2128@columbia.edu
office hours:	Tuesday, 2-4pm	tba

Course Objectives

This course will teach the fundamentals of analyzing numerical data in a social science context. Students will learn effective ways of presenting informational summaries, the use of statistical inference from samples to populations, and the linear model which forms the basis of much social science research. Emphasis will be on an intuitive understanding of statistical formulae and models, and on their practical application.

Readings

There is only one required textbook for the class:

- Moore, David S. and George P. McCabe. 2004. Introduction to the Practice of Statistics, 5th edition. New York: W.H. Freeman and Company.

This book should be available at the Columbia University bookstore.

Course Organization

This course will consist of biweekly lectures (Tuesdays and Thursdays). There will be weekly homework assignments, available from the course website, and due the Tuesdays after they are assigned.

There will be a midterm and a final examination for this class. The final exam will be cumulative, but disproportionately focused on material presented after the midterm exam.

Grading

Homework will count for 30% of the grade, the midterm will count for 30% of the grade and the final will count for 40% of the grade. The two lowest homework grades will be dropped.

Course Outline

1. Describing Data
 - (a) Administrative and philosophical overview
 - (b) The idea of a distribution
 - (c) Measures of center and spread
 - (d) Measuring relationships with categorical variables
 - (e) Measuring relationships between quantitative variables
 - (f) Interpreting OLS regression and transformations
 - (g) Multivariate regression
 - (h) A research example
 - (i) Regression diagnostics: outliers and influential points
 - (j) Regression diagnostics: aggregate data, collinearity, and causality
 - (k) Review
2. Midterm (2/28)
3. Data Collection and Statistical Inference
 - (a) Data collection (experiments and surveys)
 - (b) Probability rules
 - (c) Random variables
 - (d) The sampling distribution
 - (e) The binomial case of the sampling distribution
 - (f) Confidence intervals and hypothesis tests
 - (g) Errors in hypothesis testing: statistical significance and power

- (h) t-tests and the t-distribution
- (i) Inference for contingency tables
- (j) Inference for bivariate regression
- (k) Inference for multivariate regression
- (l) Non-parametric tests
- (m) Review