

---

## Statistics for Sociological Research I, Spring 2011, TTH 2-3:15, ASY 1101

---

<b>Instructor:</b>	Meredith A. Kleykamp, ASY4125, Phone:405-6409, <a href="mailto:kleykamp@umd.edu">kleykamp@umd.edu</a>
<b>TA:</b>	Joanna Kling, <a href="mailto:jkling@socy.umd.edu">jkling@socy.umd.edu</a>
<b>Office Hours:</b>	Kleykamp TBA, and by appointment Kling, TBA, and by appointment
<b>Textbook:</b>	<a href="#">"Introduction to the Practice of Statistics, 6th ed."</a> by Bruce A. Craig, David S. Moore, & George P. McCabe Additional readings noted on class schedule
<b>Other useful resources:</b>	<a href="#">Common Errors in Statistics</a> by Phillip Good & James Hardin 2006. <a href="#">Academic Technology Services at UCLA</a>
<b>Other requirements:</b>	Calculator with $\sqrt{x}$ , exp, and ln functions that does not store formulas Regular access to computer (I'm a PC user) Regular access to Blackboard and e-mail
<b>Prerequisites:*</b>	college algebra course undergraduate statistics course or the SOCY winterterm statistic course

### Course Overview:

This course introduces basic descriptive and inferential statistics, and various techniques of data analysis and their application in sociology. This course introduces you to both statistics and data analysis—they go hand in hand. The two basic objectives of statistical methods are data reduction and statistical inference. Data reduction involves the summarization of complex, large and multivariate datasets with an eye to the identification of patterns or associations. Statistical inference quantifies uncertainty about those summaries and associations, uncertainties that arise because data are subject to random variation. More practically in sociology, inferential statistics allow us to draw conclusions from data from random samples drawn from a population of interest.

The course is divided into a few sections. First we discuss basic tools for data reduction, description and analysis. Next we review some basic probability theory, enough to motivate and situate the connections between randomness, sampling, and inference about a whole from a part (i.e. inference). We then spend the rest of the course on estimation.

The main objectives of the course then are to:

- provide a foundation in data analysis and statistical inference
- to prepare students for more advanced graduate courses in social statistics
- to develop students' capacity to read and critically evaluate published research in the social sciences
- to begin to develop students' ability to carry out their own research using quantitative data and statistical models

---

\*You should review the math refresher on the course Blackboard site (hereafter 'Bb') and evaluate your preparedness for the basic math and algebra in this course. If you are not comfortable with these concepts you will find the math in this course challenging. You should review college algebra skills and be prepared to use those skills in this course. A college algebra course is a pre-requisite for the course. If it's been a long time since you took it—you should expect to spend ample time refreshing yourself on those skills BEFORE class begins.

## Homework

Homework will be due on most Thursdays, with due dates shown on the class schedule. There are 10 homeworks worth 10 points each. Late, missing or low scoring homework (below an 8) can be (re)submitted within 2 weeks of the original due date, and can earn a maximum of an 8.

We will rely on using statistical software for at least some portion of your homework assignments. For any homework involving the use of software, you should turn in as an appendix an edited log file that shows what you did to generate your results along with the relevant results. Relevant is the operative word here; please only include the commands and output, along with comments to explain what you are doing in those command in the edited log file. Homeworks will be graded not only on the correctness of the answer, but the clarity of your presentation.

I will instruct with Stata but you may use whatever program you prefer (SAS or R). However, we cannot help with other software programs if you choose to use them. SAS and STATA are available on the shared network drive (i:/) at most computers in the Art-Sociology Building, and at several campus computer centers, such as the OACS center in 225-227-229-231 LeFrak Hall. You will need a BSOS account; if you don't already have one, let me know. For anyone considering purchasing Stata for personal use, please consult with me before buying it, so I can advise you on which version best suits your long-term needs.

## Lectures

Lectures typically cover topics in the text and more. I will post lecture notes on Blackboard; I'll do my best to have them up before class whenever possible.

## Examinations

There will be a one-hour fifteen minute midterm during a class period (3/17, note this is the Thursday before Spring Break. Please do not make any travel plans that take you away from campus this day). The final examination will be a take-home examination, due at the time scheduled for this class by the university, in this case 5/16 at 12:30pm. You may bring a single sheet of notes and calculators that doesn't store formulas to the midterm. All work for the final exam is to be conducted independently, without consultation with fellow students.

The University of Maryland, College Park has a nationally recognized Code of Academic Integrity, administered by the Student Honor Council. This Code sets standards for academic integrity at Maryland for all undergraduate and graduate students. As a student you are responsible for upholding these standards for this course. It is very important for you to be aware of the consequences of cheating, fabrication, facilitation, and plagiarism. For more information on the Code of Academic Integrity or the Student Honor Council, please visit <http://www.shc.umd.edu>.

## Grading

40% of your grade is based on your homework: 4% each assignment for 10 assignments

30% of your grade is based on the Midterm Exam

30% of your grade is based on the Final exam

Grades will be assigned as follows:

90.0 - 100 = A or A-

80.0 - 89.9 = B+ or B or B-

70.0 - 79.9 = C+ or C or C-

60.0 - 69.9 = D+ or D or D-

0 - 59.9 = F

Plus- or Minus- grades do not affect your GPA.

If you disagree with a grade you received (on a homework or exam), please come to my office hours.

Students come to this class with a range of statistical backgrounds and aptitudes. This means that some students might work harder and improve more than others, yet still receive a lower grade. To recognize improvement, I reserve the option to slightly overweight the final and underweight the midterm for students who show improvement in their test scores.

## Incompletes

University policy allows for incompletes in unusual cases where students have completed most of the class expectations, but cannot finish the class due to unforeseen circumstances. Avoid incompletes if at all possible; students who take incompletes seldom do as well as if they had completed the course on time.

### **Class breaks**

Don't wait to formulate some "intelligent" question. When something is not clear, just ask to go over it again. Class runs from 2:00 to 3:15, which exceeds most humans' capacity for nonstop learning. We will take a three-minute break at 2:40. Be back promptly or get left behind.

### **Accommodations:**

See me in the first week of class if you have a documented disability to make arrangements for accommodation of that disability. Similarly, see me if any class-related activity, including exams, conflict with religious observances of a generally recognized nature that you are under obligation to participate in.

### **Communication with Instructor and TA:**

Office hours are times when we are guaranteed to be available (barring an emergency); we can make appointments to meet if you are unable to come to regularly scheduled office hours. Math is difficult to explain over e-mail so it's usually best to try and solve issues in person. Quick points of clarification are obviously different from fundamental questions about material, and much easier to answer efficiently over email.

### **Something to consider/unsolicited advice**

For those of you who expect to do quantitative research in your career, I would encourage you to use this course as an opportunity to learn how to either use the equation editor in MS Word, or, better yet, learn to use  $\text{\LaTeX}$ , a typesetting language that lets you deal much more easily with equations and formulas than does MS Word. It also lets you make prettier documents. There is certainly a STEEP learning curve, but if you think you may do more technical work, it may be worth learning the ropes now, using homework assignments as the testbed for writing "real" papers with  $\text{\LaTeX}$ . I'm happy to hold an optional session on getting started with  $\text{\LaTeX}$  for anyone who might be interested.

Now is the time to establish good habits in using statistical software, doing data analysis and writing up the results. Establishing a workflow early is good practice, and helps you avoid pitfalls later on. A good resource on this is "The Workflow of Data Analysis Using Stata" by J. Scott Long. It may seem like overkill right now, but it's a good idea to anticipate some of the issues you'll face as you begin working on your own research. This article is also totally worth reading: <http://www.kieranhealy.org/files/misc/workflow-apps.pdf>. Just don't be too distracted by the beautiful aesthetic quality of the document, or the emphasis on R (you can use Stata and  $\text{\LaTeX}$  easily) lest you forget the important message of the article: find and use tools that help you get work done, not create more of it.