Probability is the science of uncertainty.
— G. Grimmett

Doubt is the beginning, not the end, of wisdom.
— Anonymous, quoted in Wonnacott & Wonnacott, Introductory Statistics

Statistics are important to the study of political science in many ways. The word “statistics” originally referred to the collection of information for the state – government statistics were the first statistics. This class is thus an essential part of your graduate education, and not only for those of you who will go on to conduct quantitative research. The political philosophers among you should remember that Plato believed that that study of mathematics was an essential part of the education of the guardians. The postmodernists among you should remember that you must be able to understand an argument before you can deconstruct it. The practical graduate student who would prefer not to starve realizes that the best dissertations – and job talks – involve multiple methods, one of which is often quantitative. And all of you, no matter what type of research you expect to conduct, should know that without a basic understanding of the language and logic of probability and statistics it is difficult to comprehend even the simplest article in virtually any mainstream political science journal. As a side benefit, knowledge of statistics will make you a better-informed consumer of the myriad news stories that report the results of surveys and scientific studies from other disciplines.

My goal in this class is that each of you leaves with a knowledge of what the terminology and symbols of quantitative political science mean and with an intuitive understanding of the logic and rules that underlie those symbols and terms. You also should know how to construct a data set and do basic analyses. By the end of the class you will be ready to move on to the next course in our methods sequence, 790:633, which focuses on regression analysis and prepares you for any future methods courses you may take. For the final project, students will apply what they have learned in a paper that uses a data set related to their research interests.

Readings

Knoke, David, George W. Bohrnstedt, and Alisa Potter Mee. 2002. Statistics for Social Data Analysis, 4th edition. Itasca, Ill.: Peacock. [3rd edition is OK, although page references will be off slightly and homework will be from photocopied pages from the 4th edition]

All additional readings available on course website at [http://sakai.rutgers.edu](http://sakai.rutgers.edu)

**Grades**

Your grade will be based on weekly homework assignments, three 30-minute exams, and a final project. Attendance is required. If you miss class for an excusable reason, please let me know.

**Grade formula**

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attendance/participation</td>
<td>10%</td>
</tr>
<tr>
<td>Homework assignments:</td>
<td>15%</td>
</tr>
<tr>
<td>Exams</td>
<td>25%</td>
</tr>
<tr>
<td>Final project:</td>
<td>50%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>

**Participation**

This is mostly a lecture class, but you are encouraged to ask questions and offer suggestions throughout. Being prepared for class (that is, doing the assigned readings) is part of your participation grade. During class we also will work out problems from the book together and practice SPSS analysis in the lab. For most weeks I will assign journal articles to read for their methods content, in addition to the assigned chapters from the textbook.

**Homework**

Homework is due at midnight on WEDNESDAY, the day before class. This is so that I can see whether there are any problems that I should discuss in class on Thursday. You may upload your homework to the course Sakai site or put a hard copy in my mailbox. Late homework assignments will be penalized an entire letter grade unless the lateness is excused because of illness, family emergency, court appearance, or some equally appropriate reason. To get the most out of this class, it is important that you work out the problems yourself, run your own programs, and write your own interpretations of your output. However, this does not mean that you cannot work with others on the homework. On the contrary, I encourage group efforts on the homework as helpful to the learning process. Get help from your fellow students, but think for yourself.

**Final project**

For your final project, you will write a short paper that uses any publicly available data set to demonstrate your ability to conduct basic data analysis using SPSS and to interpret the output from that analysis. The Inter-University Consortium for Political and Social Research (ICPSR) at the University of Michigan has copies of data sets from hundreds of studies: [www.icpsr.umich.edu](http://www.icpsr.umich.edu). Many researchers also make their data sets publicly available on their own websites. Ideally this project helps teach you something about your potential dissertation research or provides some fodder for your second year paper. Details to come.

**Exams**

The exams are designed to make sure that everyone learns the vocabulary and rules necessary to be literate consumers of the statistical methods we are learning. They will emphasize definitions, identification of whether a statistical technique is appropriate for a particular type of data, and interpretation of SPSS output.
Weekly topics and reading assignments

You should complete the reading for the week before coming to class on Thursday.

**Week 1. Sept. 4. Introduction to the class and to the datasets**

How do we know what isn’t so? How can a knowledge of probability and simple statistics (and good research design – which is covered in the next semester of this course) help us avoid reaching erroneous conclusions?

**Readings:**

**Week 2. Sept. 11. Describing Variables**

Central tendency, dispersion, and other ways of summarizing variables. Also, a bit on measurement.

**Readings:**
*Phillips, Ch. 1.*
Knoke, Bohrnstedt, and Mee, Ch. 1 & Ch 2

**DUE Wednesday, Sept. 10:** KBM Ch. 1 problems 4-9. Look for a data set.

**Week 3. Sept. 18. Probability.**

The basis of all statistical inference.

**Readings:**
*Phillips, Ch. 2-4.*
Readings in Examples folder (abstract and tables only)

**Homework:** Look for a data set. Run through SPSS tutorial.
**DUE Wednesday, Sept. 17:** KBM, Ch. 2 problems 1, 2 & 5-10. ALSO: Dice spreadsheet

**Week 4. Sept. 25. Statistical inferences.**

Samuel Johnson has been quoted as saying that “You don’t have to eat the whole ox to know that it is tough.” What can we learn from a sample? Does it matter how the sample is drawn? Probability theory, Z-scores, the central limit theorem, and t-tests too. BONUS: How does non-randomness and missing data affect analysis? What can be done about it?

**Readings:**
*Phillips, Ch. 5, 7-9*
Knoke, Bohrnstedt, and Mee, Ch. 3.
King missing data readings, especially pp 49-51 & 62-65.

**DUE Wednesday, Sept. 24:** WW problems 3-9 to 3-13.
**Week 5. Oct. 2. Using SPSS and Data Set Management**

EXAM 1

No reading assignment

DUE Wednesday Oct. 1: Description of SPSS data set and plan for paper.

**Week 6. Oct. 9. ANOVA (Analysis of Variance)**

Testing hypotheses involving more than two means.

Readings:
*Phillips, Ch. 10 (section on ANOVA).*
*Knoke, Bohrnstedt, and Mee, Ch. 4*
Readings in Examples folder (abstract and tables only)

DUE Wednesday, Oct. 8: KBM Chapter 3 homework

**Week 7. Oct. 16. Analyzing categoric data.**

Also known as cross tabs. Chi-square and other tests of significance. Association between two discrete variables.

Readings:
*Phillips, Ch. 10 (Chi squared section)*
*Knoke, Bohrnstedt, and Mee, Ch. 5*
Readings in Examples folder (abstract and tables only)

DUE Wednesday, Oct. 15: T-tests and Anova using your own data sets.

**Week 10. Oct. 23. Controlling for other variables.**

Controlling for a third variable in cross-tabs, the partial correlation coefficient.

Readings:
*Phillips, Ch. 11.*
*Knoke, Bohrnstedt, and Mee, Ch. 7.*
Readings in Examples folder (abstract and tables only)


EXAM 2

No reading assignment

DUE Wednesday, Oct. 29: Human Subjects Certification.
**Week 10. Nov. 6. Association between continuous variables.**

**Readings:**
*Phillips, Ch. 6.*
Knoke, Bohrnstedt, and Mee, Ch. 6.
Readings in Examples folder (abstract and tables only)

**DUE Wed., Nov. 5: Cross tabs and controlling for a third variable using your data sets.**

---

**Week 11. Nov. 13. Multivariate regression.**
Controlling for other variables and interpreting your output.

**Readings:**
Knoke, Bohrnstedt, and Mee, Ch. 8.
Readings in Examples folder (abstract and tables only)
King and Luskin on “How Not to Lie With Regression”

---

**Week 12. Nov. 20. More on multivariate regression.**
More on regression, including when to use it and when not to.

**Readings:**
Knoke, Bohrnstedt, and Mee, Ch. 8.
Readings in Examples folder (abstract and tables only)

**DUE Wed., Nov. 19: Correlation and multivariate regression using your own data sets.**

---

**Week 13. TUESDAY Nov. 25**

EXAM 3

No homework or reading assignment.

---

**Week 14. Dec. 4. Advanced Topics.** What are probit, logit, two-stage least-squares, MLE, and other methods you commonly see in journal articles? How can you be a more knowledgeable consumer of these and other methods that you don’t yet know how to use? A taste of what is to come if you continue your quantitative training.

**Readings:**
Readings in advanced topics folder

---

**FINAL PROJECT DUE BY 5 PM, MONDAY, DEC. 15**
(One-week extensions may be available, but talk to me ahead of time.)