

Course Syllabus for Sociology 592
Advanced Social Statistics I
Fall 2004

Instructor: Richard A. Williams
741 Flanner (631-6668)
Office Hours: 9:15-10:15 MW
Email: Richard.A.Williams.5@ND.Edu
Personal Web Page: <http://www.nd.edu/~rwilliam/>
Course Web Page: <http://www.nd.edu/~rwilliam/xsoc592/index.html>

TA: Min Zhang
839 Flanner
Office Hours: TBA and by appointment
Email: mzhang@nd.edu

Time and Place: Class MW 11:45-1:00, 240 Debartolo
Lab F 1:55-3:10, 228 Debartolo

Required Texts (Available in bookstore):

Spss 12.0 Statistical Procedure Companion, by Marija J. Norusis. 2003.
Statistics with Stata, updated for Version 8. Lawrence Hamilton. 2004.
Multiple Regression: A Primer, Paul Allison (1999),

All three books will also be used in Soc 593.

I will post notes and handouts on the course web page, which you should run off and bring with you to class. The tables contained in Appendix E on the course web page should be brought with you to every class, since we will need them from time to time.

Recommended Texts: (You can borrow them from 839 Flanner)

Statistical Methods for Business and Economics, Fourth Edition, by Donald L. Harnett and Ashok K. Soni. 1991. Addison Wesley.

Statistics, Fourth Edition, by William Hays. Holt, Rinehart and Winston, 1988.

Classroom Format: The web page will contain most of my lecture notes (although I may deviate from them as the need arises). I expect you to be familiar with these notes before each class. That is, you should at least have skimmed through them even if you don't understand them. Classes will be much more productive if you are prepared to ask questions about the things you don't understand. The web page will also have the homework assignments, tables, and previous exams with answer keys. You might want to use these as "practice" exams, or simply look at them for additional examples of solved problems. (Note that organization of the course has varied a bit across time, so material on each old exam may not correspond exactly to what you will be tested on.)

SPSS & Stata: In the past, I have primarily used SPSS in this course. However, more and more faculty and students are using both SPSS and Stata, and some use Stata exclusively. Both have their strengths and weaknesses. I think it is good to be familiar with both packages, so both will be used this semester.

Readings: The SPSS and Stata texts may be especially helpful to computer novices but even advanced users should find them helpful. These books provide detailed examples of setting up problems in both programs. But, they also offer intuitive overviews of the logic behind statistical procedures, so they may help to give you a clearer picture of exactly why you are doing something. The first several chapters of each are mostly an introduction to the program and its most basic routines; I won't go over them much but I think you'll find it helpful if you at least skim through them and use them as a reference when necessary.

The optional texts each have certain strengths. Harnett's presentation is very similar to my own (partly because I stole a lot of ideas from his third edition) and he provides many examples. You may find his generally more detailed explanations extremely helpful at times. I like this book very much.

Hays is more difficult. My course notes often refer to Hays' book. You may want to refer to Hays for more in-depth discussions of some topics. Hays may be especially helpful near the end of the semester when we go over topics that are not covered in the other texts. In general, I suspect you won't find Hays very useful, but some students who want a more advanced treatment like it very much.

Many people in the past have told me that the course notes were sufficient for their purposes. Nevertheless, the readings can go into much more depth than I can in class. They also provide additional examples. If you are having trouble in class, or just want to know more about a topic, these books may be able to help you.

Examinations & Grading: There will be three open-book examinations. The best exam will count for 40% of your grade, the second best for 30%, the weakest exam will count for 20%. The remaining 10% of the final grade will come from the homework assignments. Note that under this system, a poor exam will have only a modest effect on your total grade. In the past, more than one student has started off with a C and ended up with a high grade for the course.

I usually give people two hours for exams. In the past, students have generally preferred that the exams be held on Friday or at some other mutually agreeable time. I have therefore scheduled accordingly. In exchange for taking the exam outside of regular hours, I am usually able to offer additional review time and/or cancel a class at some point during the semester (e.g. the Wednesday before Thanksgiving).

Assignments: In addition to reading assignments, there will be 10 problem sets, each of which will count for 1% of your grade. Each satisfactorily completed problem set will receive an "A", all others will receive an "F". Satisfactory completion does not require that you get every problem right. Rather, satisfactory completion requires that you make a good-faith effort to work the problems, hand them in on time, and respond to the TA's or instructor's directions for doing the homework (e.g. if the TA says you aren't showing enough of your work, you should start doing so.) Some people tend to underestimate the impact of the homework on their grade. This is unfortunate, because there are almost always at least two or three people whose final grade is different because they either did or did not do all the assignments. The homework's greatest value is that you will undoubtedly find that strong effort on the homework assignments is the biggest single thing you can do to improve your performance on the exams.

The homework assignments will also be on the web page. After assignments have been handed in, I will post the answer key on the web.

Lab: During the typical lab session, the teaching assistant will begin with a prepared presentation of routines/syntax/examples relating to the course and homework exercises. There will also be opportunities for you to ask questions about the homework and course material. The lab may occasionally include material that I do not have time to cover during regular class sessions.

Miscellaneous: You'll undoubtedly want to get yourself a scientific calculator if you don't have one already. I have seen perfectly adequate ones in the \$10-\$20 range; more expensive calculators with multiple memories and built-in statistical functions may make things a little easier for you in this course, but in the long run they will probably have more features than you need. You may also wish to take a few classes at the computing center. I strongly believe that working simple problems by hand is the best way to learn basic statistical methods, but you will also develop the elementary skills of working problems on a computer.

Also, you should activate your ND Email account ASAP. I want to be able to send email to you at this account. Even if you don't want to use your nd.edu account for email, you should activate it so you can get mail forwarded to the account you do use.

I strongly recommend that you do your computing work on a Windows machine rather than a Macintosh. If you simply must use a Mac, you may be on your own for solving any computer problems you encounter. Also, far more statistical software is available for Windows. If you use the campus WinXP machines, you should find that all the software you need for this course is readily available. If you have your own computer, I suggest getting the following software if you don't have it already:

Suggested Software for Your Personal Computers	
SPSS	ND has a site-license for this software. Contact Mary Toll at 631-5557 to get a copy. SPSS's drop-down menus make it very popular with both beginning and advanced users, and the fact that it is free for ND graduate students adds to its popularity.
Stata	Stata is an excellent program. It can do some things SPSS cannot do, and many people find it easier to work with. Alas, one thing people don't like is that it costs money if you want your own personal copy. The computer lab copies of Stata are fine if you do not want to order your own. Ordering info is at http://www.stata.com/order/new/edu/gradplans/gp3-order.html
Adobe Acrobat Reader	Available for free on the web at http://www.adobe.com/products/acrobat/readstep.html . You must have Acrobat Reader to be able to read and print the course notes. You may eventually want to splurge and buy the complete Adobe Acrobat. This software lets you create your own PDF files.
Microsoft Office 2003	Available to ND students and faculty for around \$25 or \$30 from the OIT Solutions Center. A great price. Besides Word and Excel, I suggest you also install the optional Equation Editor. Wordperfect is also ok, although I personally do not like it very much.
Others	Mary Toll also handles the distribution of SAS and SYSTAT (which are similar to SPSS) if you want them. For around \$70, she can also get you HLM. Mary Lee in LSR (631-4521) handles distribution of LIMDEP and LISREL. HLM, LIMDEP and LISREL are advanced programs that you won't need in this class but may want someday.

Tentative schedule
Sociology 592
Advanced Social Statistics I
Fall 2004

This schedule is hopefully more or less firm, but I reserve the right to make minor modifications (slightly reorder the presentation of topics, move exams up or back a little). As the number of “catchup days” suggests, it isn't unusual to be a little behind at times and then get back on schedule later. Also, the due dates for homework may be slightly adjusted depending on how we are progressing in class.

Before each class, you should read the corresponding section from your course notes. You should also read the SPSS & Stata books. You may also wish to read some of the optional material, especially from Harnett.

Date	Topic	Assignment	Readings (SPSS and Allison are required, others are optional)
8/25	Intro; Descriptive & Univariate statistics		SPSS ch. 1-5 (skim), 6 Stata 1, 2 (skim), 4 Harnett ch. 1, A.1, A.2 Hayes 2, 4.1-4.7, 4.10-4.15, Appx A
8/30	Probability		Appx D, Course Web Page Harnett, ch. 2 Hayes 1
9/1	Probability, Probability distributions		Harnett, ch. 3 Hayes Appx B
9/6	Expectations, Binomial Distribution		Harnett, ch. 4 Hayes, 3.1-3.15, 4.8, 4.16, 4.17, 6.1-6.7
9/8	Normal distribution	H # 1 Due	Harnett, ch. 5 Hayes 5.1-5.9, 5.12, 6.8-6.11, 7.16
9/13	Estimation, Confidence intervals		Harnett, ch. 6, 7 Hayes 7.1-7.2, 7.6-7.15
9/15	Intro to Hyp testing	H # 2 Due	SPSS 7 Harnett, 9.1-9.3
9/20	Review, Catch up		
9/22	Review, Catch up (If necessary)	H # 3 Due	
9/24	Exam 1		
9/27	Sampling distributions, one sample tests		SPSS 8, pp. 457-459 Stata pp. 141-146 Harnett, 9.4 Hayes 8.1-8.6, 6.4
9/29	Two sample tests, part 1		Stata pp. 146-149 Harnett, 9.5 Hayes 8.7-8.10, 8.16-8.18
10/4	Two sample tests, part 2 & 3		
10/6	Categorical Data analysis	H # 4 Due	SPSS 10

10/11 One way Anova		Harnett 5.5, ch. 16 Hayes 9.1, 9.2, 18.1-18.5, 18.15 SPSS 9 Stata pp. 149-152 Harnett 15.1-15.5 Hayes 10.3-10.20,12.1-12.8,10.23
10/13 Two way Anova	H # 5 Due	Stata pp. 152-154 Harnett 15.6-15.8 Hayes 12.17
10/25 Choosing the Right Procedure: Examples		
10/27 Review, Catch up	H # 6 Due	
11/5 Exam 2 (The later date is because people usually want a little more time after the break.)	H # 7 Due on 11/3	
11/1 Bivariate regression, Part I & II		SPSS 11, 12 Stata 6 Allison 1, 2, 4, 5 Harnett ch. 11, 12 Hayes 14.1-14.3,14.9,14.11,14.6
11/3 Bivariate Regression, Part II		Hayes 14.13-14.17
11/8 Multiple Regression - Intro		SPSS, 23 Harnett, ch. 13.1-13.5, 13.7 Hayes 15.8,15.9,15.13-15.15
11/10 Multiple Regression - Intro	H # 8 Due	SPSS 13
11/15 MR - matrices, standard errors, multicollinearity		Allison 7 Harnett, 13.8 Hayes 15.21, 15.26
11/17 MR - Standardized coefficients	H # 9 Due	Hayes 14.4, 14.8
11/22 Qual IVs; F tests for IV subsets; Anova and Regression		Harnett 13.10, 15.9 Hayes 15.2, 15.4, 15.6, 15.10, 15.11
11/24 NO CLASS (Compensation for taking exams on Fridays)		
11/29 MR - Semipartial & Partial coefficients; Stepwise MR		Harnett 13.11 Hayes 15.17, 15.19, 10.1, 10.2, 16.1
12/1 Violations of assumptions; Analyzing residuals (Topic subject to change)		Stata 7 Harnett 13.6, 13.9
12/6 Review, Catch up (If necessary)		
12/8 Review, Catch up (If necessary)	H # 10 Due	
12/15 Exam 3, 8:00-10:00 am (Subject to change to a mutually agreed upon time)		
