Socy 10: Quantitative Analysis of Social Data

Summer 2009

Matissa Hollister
Silsby 104

Every day we hear statements like the following:

Americans are more worried than ever about the economy.
Teens in families that eat dinner together have a lower risk of drug use.
Schools with smaller class sizes have better outcomes.
Youth who grow up in good neighborhoods are more likely to go to college.

How would you test these hypotheses using quantitative data? What are the appropriate techniques to use for the analysis? If someone else has conducted a study, how can you tell whether to trust the results?

This course provides you with the basic tools to be both informed consumers and responsible creators of social science data. The emphasis of the class is on understanding the concepts and appropriate uses of quantitative methods rather than deriving and memorizing formulas. That being said, the class does involve a considerable amount of calculating and working with numbers. The key to success for this course is attending class, keeping up with the material, and asking for help when you're confused.

Textbook

I do not require you to purchase a textbook for the class. I have, however, selected an optional textbook for the course. This textbook may be useful for students who learn better by reading or for providing an alternative explanation of topics you find confusing. In addition, the exams are open book, so the textbook may come in handy there as well. I will put a copy of the textbook on reserve at the library. I also have copies of several different statistics textbooks in my office that I'm willing to lend out for a week at a time.

One of the reasons why I do not require a textbook is that I've yet to find the perfect book. The optional book I've selected is on the simpler side of the textbook spectrum, with the assumption that many people turning to the textbook are confused and looking for help. If you are looking for a slightly more advanced and mathematical textbook, you might look for *Statistical Methods for the Social Sciences* also by Agresti. The optional textbook for the class is:

Alan Agresti (Author), Christine A. Franklin (Author) (2006). *Statistics: The Art and Science of Learning From Data.* Prentice Hall.

Any of the following versions are fine:
First edition, text only version, ISBN 0007560710
First edition with a CD (which we won't use), ISBN 0130083690
Second edition, ISBN 0135131995

The course will cover a handful of topics that are not in this textbook. If you would like further materials to read on a topic please contact me and I will do my best to find and copy a section from one of my many other stats textbooks.

X-hours
Please note that there are several class sessions that are moved to x-hours. See also note below on review sessions.

Getting help: review sessions, blackboard discussion board, office hours, etc.
I want to help students who are struggling in the class, but I also want to avoid helping students one-by-one, answering the same question over and over again. So, I've set up a number of ways to provide help in the class in a more collective fashion:

Review sessions vs. office hours
I will hold a review session every week for students who are having difficulty with specific problems or topics in the class. This review session will usually be held during the x-hour except for the weeks when the x-hour will be used for regular class sessions. I will conduct a poll of the class in the first week to determine an alternate time for review sessions during those weeks.

Please come to these review sessions if you have questions about problem sets or material covered in class. Other students who come to the review sessions will benefit from hearing my responses to your questions. Please *don't* come to my office hours to ask these types of questions. Please use my office hours to discuss more individual or personal questions, or to meet with me about your group projects.

My office hours this summer will be Wednesdays 11:15-12:00, 1:00-2:00. I have an online signup system for scheduling office hours. To sign up, go to the following website, click on an slot, and enter your name:

http://www.supersaas.com/schedule/HollisterDartmouth/OfficeHoursF08

Blackboard discussion board vs. email
I have created a forum on the blackboard discussion board where you can post questions about the problem sets, etc. Again the idea here is that other people will benefit from reading responses to your questions. Please *do not* email me individually with questions about the problem sets.* Be assured that I am subscribed to this forum so I am notified each time someone posts on the board. I will respond to posts *at least once a day, but not always immediately*. Please feel free to respond to each other's posts. You can subscribe to the forum if you wish to be notified each time someone posts. Just click on the "subscribe" button.

*Exceptions to my "no email" rule: blackboard automatically grades the problem sets and sometimes this causes problems. If, for instance, you forgot to put in a decimal point but you clearly got the right answer, you can send me an email and I'll adjust the grade. In these types of circumstances it's obviously better to email me directly than to post a message on the class discussion board.

Assignments
Problem sets
The problem sets for the class are on blackboard. The goal of these problem sets is to give you a chance to practice and develop your knowledge of the material. With this goal in mind, you can get full credit even if you get a couple of questions wrong. Here's how it will work using the example of a problem set with 50 possible points:

- I will take 85% of the total points. In this example, 50*0.85=42.5
- The denominator for the grade will be calculated using this lower threshold:
    - If you got 42.5 or more points, you'll get 100%
    - If you got 37 points, you'll get 37/42.5=87%

Please note that blackboard will automatically grade your problem sets in terms of the points you get, but it's not capable of following such a complicated grading scheme. You can use the points reported on blackboard to calculate your actual grades using the instructions above. I will also periodically calculate the results in Excel and upload them to Blackboard. Your lowest three problem set grades will be dropped. You may work on your problem sets in groups, but each student needs to submit his/her own response on blackboard.

There will be problem sets due before each class (except exam days). They must be submitted online by 9:30am. Blackboard will control the availability of the problem sets (it will turn them off exactly at 9:30am) and late submissions will not be possible.

Blackboard automatically grades your problem set, so you get instant feedback. Some of the problem sets include ungraded essay or class discussion questions. The ungraded essay questions are designed to help you think about the topic in more depth and are examples of types of questions you might be asked on the exams. As the name implies, these questions are assigned zero points and are not graded by Blackboard. Blackboard will, however, provide the answers to these questions so you can evaluate your own response. The class discussion questions are also ungraded and are questions I'd like you to think about and be prepared to discuss in class. I will randomly select students from the class to start discussions on these questions.

Exams
There will be three exams. Each exam will include an in-class portion as well as a take-home essay section. You may use your textbook, class notes, and any other written materials for the exams. The only electronic device for the in-class sections should be a calculator.

The take-home essay questions will be handed out during the exam and must be completed by 5pm two days later. Your answers to the essay questions should be written on a computer and submitted electronically. Again, you can use your textbook, class notes, and written materials. However, you may not discuss the problems or your answers with anyone else (students in the class, friends, or anonymous strangers), in person or electronically (including email, online discussion boards, etc.). These essay questions are not designed to take you the whole evening. They are intended to take approximately an hour of your time and will have maximum page limits to restrain over-eager essay writers.

Class Project
An important aspect of the class is the group project. See the attached description for more details. You'll do the projects in groups of 3-4 students on a topic of your own choosing. These projects will require significant time and commitment, so choose a topic you're enthusiastic about and get started early! At the end of the term, you will be asked to evaluate the contributions of your group's members, including yourself, to the project. These evaluations will affect your project grade.

Grading
Grades for each part of the course will be weighted in the following manner:

Problem sets: 10%
Exam 1: 20%
Exam 2: 25%
Exam 3: 25%
Group project: 20%

Grades are set according to the following scale. (*Note that I do not round up!)*:

100-93: A
92.99-90: A-
89.99-87: B+
86.99-83: B
82.99-80: B-
79.99-77: C+
76.99-73: C
Etc.

Missed classes
I do not take attendance in this class, but you will find it challenging to do well if you are not present for the lectures. I am not responsible for helping students catch up on material they missed. It is your responsibility to get the appropriate information and materials from classmates.

Please let me know at least 10 days in advance if you will be unable to be in class for one of the exams. If you are sick that day or have some other emergency, you will be asked to provide documentation (from Dick's house, deans, etc).

Electronic devices
Calculator
You should have a calculator and bring it to class and exams. It doesn't have to be fancy, but it should be capable of squaring and taking the square root.

Laptops
Computers are fantastic tools for statistical analysis, but poor tools for taking notes in class. The availability of wireless also makes surfing the web an irresistible temptation. For these reasons, **you should not bring your laptop to class.** If you rely on your laptop to take notes, you may seek permission to bring it to class. Be prepared to show me how you plan to write out formulas and draw graphs on your computer.

Cell phones
Please turn off your cell phones while in class. I reserve the right to answer any phone that rings in class.

Statistical software
We'll be using the statistical software program Stata for problem sets and your final project. Stata is available for free on the Dartmouth network for both Macs and PCs. It is a key-served program, though, which means you have to be connected to the internet and either on campus or running VPN in order to use it.

The following website contains information on Stata resources for students:

http://www.dartmouth.edu/comp/support/library/software/statistics/students.html

And the following website provides instructions on how to install Stata on your computer:

http://www.dartmouth.edu/comp/support/library/software/statistics/stataintro.html

You will need to have Stata up and running on your computer by the second week of class. If you have problems installing Stata I'm not going to be much help, please use the resources above and the computer help desks.

Most of the public computers around campus, including the library computers, have Stata already installed on them. Therefore, if you have problems installing the program or your laptop is in the shop please use one of these computers. A problem with your laptop is not a valid excuse for not turning in your problem sets.

Academic integrity

I expect students in the class to conduct themselves in accordance with Dartmouth's honor code and with academic and personal integrity. Explanations of Dartmouth's integrity rules and principles can be found at http://www.dartmouth.edu/˜uja. Students are expected to take responsibility for doing their own work, providing proper citations whenever using words or ideas borrowed from others. Details on citing sources are available at http://www.dartmouth.edu/˜sources. I also expect you to respect other students while in class and when working together on projects.

Disabilities & religious observances

Students with learning, physical, or psychiatric disabilities enrolled in this course that may need disability-related classroom accommodations are encouraged to make an office appointment to see me before the end of the second week of the term. All discussions will remain confidential, although the Student Accessibility Services office may be consulted to discuss appropriate implementation of any accommodation requested.

Some students may wish to take part in religious observances that fall during this academic term. Should you have a religious observance that conflicts with your participation in the course, please come speak with me before the end of the second week of the term to discuss appropriate accommodations.

Course schedule: topics may be shifted as necessary.

Day	Topic	Textbook Chapters	Lecture #	PS # Due	Project Due Dates
Friday	Intro/overview of course				
Monday	Formulating research questions/sampling	Ch. 4	1		
Wednesday	Types of variables, frequency distributions	2.1	2	1	
Thursday (X)	Intro to Stata		3		
Friday	**College holiday- No class**				
Monday	Graphs, distributions & more on Stata	2.2	4	2 & 3	
Wednesday	Measures of central tendency & variability	2.3 & 2.4	5	4	Presentation of ideas
Thursday (X)	Cross-tabs	3.1 & 3.1	6		
Friday	Scatterplots & correlation	3.1 & 3.2	7	5	
Monday	**No class**			6	
Wednesday	**No class**			7	
Friday	**No class**				Proposals due
Sunday	*8:30pm- Optional exam review session*				
Monday	**Exam on lectures 1-7**				
Wednesday	Scatterplots, correlation & regression	3.2,3.3,11.1, 11.2	8		
Friday	Regression with dummy variables, cautions with regression	3.4	9	8	
Monday	Intro to multiple regression	12.1	10	9	
Wednesday	Omitted variable bias		11	10	
Thursday (X)	Omitted variable bias (cont)		12		
Friday	Nested models		13	11 & 12	
Monday	Review session			13	
Wednesday	**Exam on lectures 8-13**				
Friday	Sampling distributions	Ch. 6	14		Progress meetings
Monday	**No class**				
Wednesday	Confidence intervals	Ch. 7	15	14	
Thursday (X)	Confidence intervals	Ch. 7	16		
Friday	Hypothesis testing	Ch. 8	17	15 & 16	
Monday	Hypothesis testing	Ch. 9	18	17	
Wednesday	Hypothesis testing	Ch. 9	19	18	
Friday	Chi-square	Ch. 10	20	19	
Monday	Inference in regression	Ch. 9, 11.3	21	20	
Wednesday	Project poster session				Poster session
Saturday	Project report due **5pm**			21	Report due
Monday	**Exam on lectures 14-21, 8:00am**				