Introduction to Quantitative Methods in Geography
GEOG 4023/5023

Instructor	TA
Nicholas Nagle	Lisa Jordan
Phone: 303-492-4794	303-492-6062 (No messages)
Office: Guggenheim 207	Guggenheim 301
Office Hours: Wed. 2-3	Wed. 10-10.50 Gugg 301
Thu. 2-3	Fri. 10-10.50 KESDA
or by appointment	or by appointment
email: nicholas.nagle@colorado.edu	Lisa.Anselmi@colorado.edu

wepage: www.colorado.edu/geography/courses/geog_4033_f05

Some (hopefully calming) words regarding the text

When many of you bought the textbook, I am sure that you looked at it, saw some calculus, and either gasped, cursed my name, or did both. Let me first state that this is not strictly a statistics course, it is a quantitative methods course. Calculus is how we quantify things that change. Unless you believe in a world that doesn't ever change, you should be slightly familiar with calculus. Calculus is not a prerequisite of the course – I will cover the basic elements. But of course that doesn't make you any less scared, now does it? So here are a few points that I hope will make it less scary.

1. Failure to understand the nitty gritty details of calculus *might* mean the difference between an A and an A-. No more. I promise. Maybe not even that. All I ask is that you try.
2. The most important important things that I expect you to remember about calculus are at a very basic conceptual level that can be understood by drawing a curve. I think that the concepts of calculus are very important in any quantitative science. The actual *doing* of calculus much less so. In fact, you already use calculus daily at this conceptual level without realizing it. My goal isn't that you actually know how to do calculus, but that when you see it, you will know what is happening, without necessarily knowing how to do it yourself. Hopefully I can demystify it.
3. The level of calculus that is required to fully grasp the why of statistics is less advanced than the level of treatment calculus it is given in high schools.
4. The calculus you will need to know can be looked up in tables I will provide you.
5. I will hold your hand through this. I am not making you do this to improve my health; in fact, I expect my own health to decline because of this. This is a lot of work on my part also, but this is the way that I think it should be done.

The most difficult portions of the book that I will assign are Chapters 2 and 3. We will spend around 5 lectures in these chapters. After that, you will see practically zero integrals. Promise.

Prerequisites

I do assume that all of you have basic abilities with algebra, however, there are no formal prerequisites for the class. I will cover concepts from calculus and linear algebra as needed.

Audience

Graduate and advanced undergraduate students primarily in geography. Examples will be drawn from atmospheric science, hydrology, political geography, economic geography and other disciplines of human and physical geography.

Objectives

This course is designed as a first, and for many only, course in statistical methods for geographers. Before I

begin with the objectives, let me explain what I think constitutes good statistical analysis. Before statistical analysis can begin, you, as a scientist, (and not as a statistician) must have a solid understanding of your research question, and from there, an explicit statement of a testable hypothesis. This is where the statistical analysis begins. We must then choose a method that is appropriate for our hypothesis. Most methods can be "cookbooked," that is they can be applied in almost too easy a fashion, so that numbers can be spit out of the computer. At this moment, this is all that they are, numbers, and nothing more.

The interpretation of these numbers, however, requires you to be both a bit of a statistician, and lot scientist. The little bit of statistician needs to know under what conditions (i.e. under what assumptions) these numbers are exact. These numbers are only exactly correct under very specific conditions, and it is your job to know what these conditions are. The scientist in you needs to know why these conditions are most likely unrealistic or inadequate. Put differently, the statistician in you needs to know what the "perfect world" would look like, and the scientist needs to know the difference between the perfect world and the real world.

To summarize: Good statistical analysis demonstrates that the researcher knows exactly what the conditions are under which their results are correct, demonstrates the knowledge that these conditions might not all be met in reality, and provides an explanation for how these results might be interpreted under these imperfect conditions.

So what are the objectives of this course? First and foremost, I hope that you learn how to critically assess statistical analysis. At the end of the course, you should be able to identify the assumptions under which the researcher's results are exactly correct. Many times, it isn't even clear if the researcher knows this themself. If you can identify the underlying assumptions, you can then, as an expert scientist, and not as a statistician, decide for yourself what you think about their stated results.

Secondly, I hope that by the end of the course you will have enough experience to practice "good statistics" yourself. When you use a method, you will be aware of the requisite assumptions under which the results are correct. In this way, you will be able to know exactly what your results tell you, and what the possible shortcomings of your method are.

Finally, and least importantly, I hope that you will have gained the mathematical experience to understand for yourself many of the new quantitative methods that are constantly being introduced into the analysis of geographic data. Many of these methods have not yet been adopted into statistical software, and implementing them requires a certain amount of "do it yourself." The mathematics required to understand the majority of these methods, however, isn't that difficult. We will cover the mathematics and then practice, practice, practice on statistical problems. These really are simple mathematical tools compared to the advanced mathematical techniques used in many other fields. If you are comfortable with just a few simple mathematical tools, then you should be able to understand the vast majority of statistical techniques that you will ever come across.

General Notes on Course Structure

I will tell you before each lecture what material is to be read before we meet again. I expect you to have read this material before class. The common format of lecturing was designed for a time when textbooks were prohibitively expensive, so as to place them beyond the economic means of most students. Thus, the only way for the instructor to pass on knowledge was through reciting this material orally. While you may still find textbooks to be expensive, the price has indeed declined over the last millennium of higher education. Thus, I do not think that repeating the text material, which you can read for yourself, to be the best use of our time. I will try as much as possible not to duplicate the material from the text.

Mathematics plays an important role in statistical analysis. Statistical analysis is only exactly correct under a specific set of conditions, and these conditions are expressed mathematically. You need to be able to understand these before you can assess if these conditions are realistic. In addition, another objective of

the course is to develop the mathematical experience to understand new methods. To facilitate these ends, I have chosen an introductory statistical text that places a heavy emphasis on mathematics. When you read the book, it should be with paper and pencil in hand. There are many examples in the text. GO THROUGH THEM YOURSELVES. Having said that, there are many proofs in the book. I give you the liberty to skip them. Unless you plan on taking an active role in developing new statistical methods, you don't need to mathematically prove anything.

Since I am going try hard not to duplicate the text material, what will I do with lecture time? I will try to cover deficiencies in the text. I find that intuition is hard to glean from most textbooks. Thus, at some times during lecture, I will back off from the mathematics and discuss the intuition behind things. Another thing that I think textbooks do inadequately is present worked exercises. Compare the number of worked exercises with the number of end of chapter problems in most texts. This ratio should be reversed. How can I expect you to do a problem when you haven't had sufficient practice with similar problems? Thus, I will try to cover as many examples as possible in lecture time. Some examples will from the text, some from elsewhere.

Textbooks
Required:

- Rice, John A. 1995. Mathematical Statistics and Data Analysis. Duxbury Press.
  This book will be the primary reference for the class. It was written for advanced undergraduates/graduates in the maths and sciences. I expect many of you will have difficulty understanding portions of this book on the first read. Nevertheless, it is usually a good book. Unlike other mathematical statistics texts, this one is solidly focused on data analysis. We will skip all of the more challenging mathematical aspects. A note on price... as you probably noticed, this book is outrageously expensive. I found copies of earlier editions for a little over half the price at alibris.com.

- Rogerson, Peter A. 2001. Statistical Methods for Geography. Sage.
  This book is much simpler and I hope that all of you will be able to understand this book without too much difficulty.

Recommended (especially if the Rice is too much like Greek)

- Freedman, David., Roger Pisani, and Roger Purves. 1998. Statistics. Norton.
  If the Rice is Greek to you, I suggest that you see if you can find this one. Someone is selling a bunch of 1st editions at alibris.com for 3$! I have 1 personal copy you can look at, and will place another copy on reserve in the library. If any of you go this route, please, let me know how it is working out!

Grading
Homework $\times 6$: 10% each
Midterm $\times 2$: 15% each
Final Project:10%

A note on grade assignment for undergraduates:

I will hold you to a different scale than the graduate students. Here is how grade assignment most often works: at the end of the course, the instructor looks at the distribution of grades, sets the middle to a B- (or whatever grade inflation is at these days), and works out the highs and lows from there. Since there are only a few of you, this isn't an option. Doing so would only create a sense of competition when I want you to help each other out. I can promise you that if you calculate your grade based on the above percentages, (with 90% an A-, 80% a B-, etc.) that you will get at least that grade and no lower, and most likely higher.

My hope is that you will all get A's and B's.
A note on grade assignment for graduates:

Haven't you yet learned that grades no longer matter! At this point, your motivation should be intrinsic, not extrinsic.

Homework

You are allowed to collaborate with other students on the homework. In fact, I strongly recommend that you set up study groups to do so, however your assignments must be written by yourself only. I will not create groups for you, but I do ask that you be good citizens, and include members of varying mathematical/statistical background. Help each other out and everyone will benefit. Homeworks will be due at the beginning of class (no exceptions) on the stated due date. These dates are tentatively on the Friday of Weeks 3, 5, 7, 10, 12 and 15.

Midterms

Midterms will be 50 minute long. Problems will be similar to the homework problems, only slightly easier for the most part. The second Midterm will not be cumulative. The dates of the midterms are, 10/12 and 12/2. Reasons for excusable absence include religious observance, a sudden change in your own health status which requires immediate hospitalization, death in the immediate family, or otherwise at the instructor's discretion. Make-up exams may contain take-home and/or oral components. My suggestion is that you do not learn more about make-up exams.

Final Project

By **Noon, Wednesday, Dec 14**, you are required to turn in to me a "referee" report for a published article of your choosing. The article must not be by a member of the CU faculty. The article must be contain a sufficient element of statistical analysis. It need not present a novel method of analysis, but statistical analysis must be a significant element of the article. You must send the article to me on or before November 23 to get it approved. A good source of articles is any issue of *Professional Geographer*, but please feel free to choose any article you like from any journal. (Note, the above link only works from campus computers.)

Your write-up should be part report, summarizing what the author's problem is, how they set up a method, what they did, and the conclusions they arrived at. You may get help from myself or the TA in aspects of understanding what they did. We will do our best to help you identify the connections between the article and what you are learning in class, but will offer little more than that.

Most importantly, however, you must critically evaluate their method and results and explain whether you are convinced of their conclusions and why. If applicable (i.e. unless the analysis is perfect!), comment on how the research or analysis might be improved upon.

University Policies

If you qualify for accommodations because of a disability please submit a letter to me from Disability Services in a timely manner so that your needs may be addressed. Disability Services determines accommodations based on documented disabilities (303-492-8671, Willard 322, www.colorado.edu/disabilityservices).

I will make reasonable accomodations for students who have conflicts between religious observance dates and course examinations or assignments. Please contact me at the beginning of the semester if you think that you may require such accomodation. For university policies on this, see www.colorado.edu/policies/fac

Course Schedule

As with all schedules, this one is meant to be broken. Under the readings column, Rice, rather obviously, denotes relevant chapters from Rice. Since I will, however, be skipping many portions of the book, I will make it more clear ahead of time exactly which pages I want you to read. Rog. denotes readings from Rogerson. FPP denotes optional readings from Freedman, Pisani and Purves.

Week	Dates	Topics	Reading
1	8/22-8/26	Introduction, Calculus Refresher, and Probability	Rog. 1, Rice 1, FPP 13
2	8/29-9/2	Probability Distributions and Random Variables	Rice 2, Rog. 2, & page 225
		9/5 - Labor Day - No class	
3	9/7-9/9	Joint Distributions	Rice 3
4	9/12-9/16	Expectation	Rog. App. B, Rice 4
5	9/19-9/23	Law of Large Numbers, Central Limit Theorem and Sampling Distributions	Rice 5,6
6	9/26-9/28	Summarizing Data, Histograms, and Density Estimation	Rice 10.1-10.3
7	10/3-10/7	Point Estimation (Method of Moments and Maximum Likelihood)	Rice 8
8	10/10-10/12	Slack Day and Mid-Term	
		10/14 - Fall Break - No class	
9	10/17-10/21	Interval Estimation	Rice 9, FPP 26
10	10/14-10/28	T-test and Experimental Design	Rice 11, Rog. 3, FPP 26,1,2
11	10/31-11/4	Bivariate and Multivariate Regression	Rice 14, Rog. 6, FPP 10,12
12	11/7-11/11	More Regression - Testing and Residual Diagnostics	Rog. 6
13	11/14-11/18	More Regression - Problems	Rog. 7
14	11/21-11/23	More Regression - Some more problems in space and time	Rog. 9
		11/25 - Thanksgiving Break - No class	
15	11/28-12/2	Geostatistics	
16	12/5-12/7	Slack Day and Mid-Term	

Other Useful Texts Worth Looking At or Having on your Shelf in the Future

- Ramsay, Fred L. and Daniel W. Schafer. (2002). The Statistical Sleuth. Duxbury Press
  This book is excellent at discussing how the proper interpretation depends on the combination of experimental design issues and the choice of statistical testing method. It is organized around many helpful case studies and shows how analysis of these cases should proceed and how the results should be presented. It does not, however, cover any probability theory.

- Miller, Jane E. (2005). The Chicago Guide to Writing about Multivariate Analysis. University of Chicago Press.
  Everyone that includes statistical analysis into their writing should take a look at this. Covers effective ways to convey statistical data and arguments to various audiences, as well as how to interpret results from common procedures.

Subject Specific Texts
I understand that many (most?) of you will not continue with statistical analysis in your own research, but if you do, you really ought to consider purchasing a text/reference that is more particular to the practice in your chosen field. Here are a couple of good suggestions.

- Haan, Charles T. (2002). Statistical Methods in Hydrology. Iowa State University Press. 2nd ed.
  This updated classic gives many reasons for recommending it. It offers a clear presentation spanning topics from elementary probability to time series analysis and geostatistics. Lots of worked examples.

- Wilks, Daniel S. Statistical Methods in the Atmospheric Sciences. (1995). Academic Press.
  This book has an excellent discussion of applied probability models and contains a good section on multivariate techniques such as discriminant and principal component analysis. Regression is practically ignored, however.

- Cressie, Noel A.C. (1993). Statistics for Spatial Data. Wiley.
  This book is the bible for spatial statisticians. Covers geostatistics, autoregression, point patterns and more. I recommend it to anyone wishing to continue with spatial statistics after this course. It is fairly advanced, but I hope that it will be more approachable by the end of this course.

- Isaaks, E. H. and R. M. Srivastava. (1989). Oxford.
  Applied Geostatatistics. Good introductory text for geostatistics.