Econ 7818: Mathematical statistics for economists

Edward Morey
Fall 2006 (rough working draft: August 28, 2006)

Econ 7818 is the first course for Ph.D. students in Economics in the statistics/econometrics sequence. Statistics is the foundation of econometrics.

What is statistics?

This will be the topic of our first lecture. Put simply: A course in statistics studies statistics, and every function of random variables is a statistic. Econometricans develop and use statistics, then use these statistics, along with data, to estimate things about the world of interest to economists.

What is econometics?

Quoting from *A Guide to Econometrics* by Peter Kennedy, who is paraphrasing a famous play (by Oscar Wilde, I think)

> "You haven't told me yet," said Lady Nuttal, "what it is your fiancé does for a living."
>
> "He's an econometrician." replied Lamia, with an annoying sense of being on the defensive.
>
> Lady Nuttal was obviously taken aback. It had not occurred to her that econometricians entered into normal social relationships. The species, she would have surmised, was perpetuated in some collateral manner, like mules.
>
> "But Aunt Sara, it's a very interesting profession," said Lamia warmly.
>
> "I don't doubt it," said her aunt, who obviously doubted it very much. "To express anything important in mere figures is so plainly impossible that there must be endless scope for well-paid advice on how to do it. But don"t you think that life with an econometrician would be rather, shall we say, humdrum?"
>
> Lamia was silent. She felt reluctant to discuss the surprising depth

of emotional possibility which she had discovered below Edgar's numerical veneer.

"It's not the figures themselves," she said finally, "it's what you do with them that matter."

Some other quotes from *A Guide to Econometrics*:

Econometrics is what econometricians do.

Econometrics is the study of the application of statistical methods to the analysis of economic phenomena.

What distinguishes an econometrician from a statistician is the former's preoccupation with problems caused by violations of statistician's standards assumptions; owing to the nature of economic relationships and the lack of controlled experimentation, these assumptions are seldom met.

Econometricians are often accused of using sledgehammers to crack open peanuts while turning a blind eye to data deficiencies and the many questionable assumptions required for the successful application of these many techniques.

Econometric theory is like an exquisitely balanced French recipe, spelling out precisely with how many turns to mix the sauce, how many carats of spice to add, and for how many milliseconds to bake the mixture at exactly 474 degrees of temperature. But when the statistical cook turns to raw materials, he finds that hearts of cactus fruit are unavailable, so he substitutes chunks of cantaloupe; where the recipe calls for vermicelli he used shredded wheat; and he substitutes green garment die for curry, ping-pong balls for turtle's eggs, and for Chalifougnac vintage 1883, a can of turpentine. (Valavanis)

It is the preparation skill of the econometric chef that catches the professional eye, not the quality of the raw materials in the meal, or the effort that went into procuring them (Griliches)

The art of the econometrician consists in finding the set of assumptions which are both sufficiently specific and sufficiently realistic to allow him to take the best possible advantage of the data available to him (Malinvaud)

The applied econometrician: The applied econometrician, unlike the theoretical econometrician, needs to worry as much about her data as about the theory. The forecasts and predictions generated by the econometric model are only as good as the data that produced them.

Course Description:

Econ 7818 is a course in statistics for Ph.D. students in economics. Important components are probability theory, distribution theory, statistics, estimators and estimates, properties of estimators, sampling, inference, and particular estimators such maximum likelihood and OLS. Extensive use will be made of mathematical and statistical software.

Web page:

My web site is located at http://www.colorado.edu/Economics/morey/index.html From it you can link to the web page for Econ 7818, or you can go directly to web page for the course at http://www.colorado.edu/Economics/morey/7818/7818home.html .

All assignments, review questions, additional readings, lecture notes, etc. will be

made available at this site on an as-need basis.

Text and readings:
The text for this course is Alexander Mood, Franklin Graybill, and Duane Boes, Introduction to the Theory of Statistics, McGraw Hill, 1974. ISBN 0-07-042864-6.

While you want to acquire your own copy of the book. I have put it up on the course web page. The book is no longer in print, but used copies are available on the web.

The text for 7828 is *A Course in Econometrics*, by Arthur S. Goldberger. Harvard University Press, 1991. I recommend you get it now rather than waiting.

Other texts that I might have chosen for this course include:
Takkeshi Amemiya, Introduction to statistics and econometrics, Harvard University

Press, 1994, ISSBN 0-674-46225-4

For those who are considering additional books, I recommend:
Peter Kennedy, A Guide to Econometrics (4th edition), MIT Press, 1998.
ISBN 0-262- 61140-6. This is an excellent book that provides, in words, the big picture. I recommend it highly. It will help with this course, and help even more in 7818

Russell Davidson and James MacKinnon, *Estimation and Inference in Econometrics*, Oxford University Press, 1993. I recommend this book to those who want an advanced text in theoretical econometrics. It is well written. Russell and I went to graduate school together. This book is not for the faint of heart.

Additional reading and notes:
I will assign some additional readings for some topics. I will try to put these on the course web page.

Software:

The computer software *Mathematica* will be an important tool. We will use it to investigate distributions such as the Normal, Chi-Square, Student t and F distribution: distributions that play crucial roles in econometrics. We will use it to draw random samples from these and other distributions. We will use it to write our own code for different econometric estimates, to do simulations, and to do Monte Carlo studies.

The graphical capabilities of *Mathematica* will allow us to visualize important concepts.

Mathematica is available on the computers in the computer room for Econ graduate students. There is no need to purchase *Mathematica*.

While it is not necessary to purchase *Mathematica* for the course, one can purchase a PC version.

(The following information is from Bruce Fast, the CU software license guy)

> The CU-Boulder site license for *Mathematica* is for campus computers, departmental computers, staff and faculty – not for students, as such. But any grad student with some kind of monthly appointment to a department counts as "staff" for this purpose. The annual license fee under the site license is $108, though for fall semester we're charging half that for a 2006 *Mathematica* license; $54. With the fee, a staff/faculty person can license both their office/desktop computer and home/laptop computer.
>
> Students have a different way to go; they can get their own license for *Mathematica* online, as per the website: http://spot.colorado.edu/~sitelic/mathematica/studentlic.html. A 6-month version runs $45; a 12-month version, $70; a non-time-limited edition of *Mathematica* is $140. At the bookstore (CU Bookstore's "Buffalo Chip") students can get a "permanent" license plus installation CD for $150.) The student license gets locked to a particular computer, and requires some special procedure to get transferred to a different computer.

To pursue the site license option (only if you have an appointment) contact Bruce

Fast at sitelic@colorado.edu

I have a problem with *Mathematica*: I do not use it . The mathematical software I use on a day to day basis (MuPAD) is not freely available to you. So, I will not be able to answer your technical questions about *Mathematica*. I will be learning it along with you, but hopefully I will not have to learn as much as you will have to learn.

Great detail about *Mathematica* can be acquired from their web site: http;//www.worfram.com/ . See also their MathWorld site at http://mathworld.wolfram.com/ , particularly the section Probability and Statistics. You will likely want to use the MathWorld site to learn about mathematical and statistical concepts, even if you never were to use *Mathematica*. My lectures notes will sometime plagiarize from MathWorld.

On the course web page, there is a set of readings, links, and notes on *Mathematica*.

Prerequisites:

A sufficient condition for being in this course is that you are a new Ph.D. student in economics here at C.U. That said, you need to get by and do well in the course whether your preparation for it is "not enough", "just right" or "too much". I won't use matrix algebra in this course, but I am sure it will be used in your upcoming econometrics courses, so you might want to sleep with a matrix algebra book.

But, being a new Ph.D. student is not necessary to be in 7818 - you might, for

example, be a graduate student in the Business School. Or, someone with just too much free time.

Class format:

Lecture/problem solving/discussion/computer/estimation

I very much believe that one learns statistics by doing, rather by listening to me

drone on.

Hands on experience will be stressed. Class format will include both individual and group problem solving. We will extensively use random sampling to investigate the properties of a statistic, and then relate what we find to statistical theory. After completing the course, you will be better able to critically evaluate and apply econometric theory.

You will spend a considerable amount of class time interactively formulating and solving problems and building models. Small groups will often be utilized.

I am hoping that you will spend more time talking than I will spend talking.

Details of grading:

There will be problem sets, short projects and quizzes. Some of these activities will be done using *Mathematica*. Some of these activities will be done in groups larger than one. Some of these activities will be take-home, some will be done in class. Each will be graded on a 10 point scale. Your best $(N-2)$ grades on these activities will constitute 50% of your course grade, the midterm 20%, the final 25%, and 5% will be set aside for participation (I will be the "decider" when it comes to how much you participated). The final will be cumulative and might have two parts: a take home part and an inclass part.

Wrt the problem sets, short projects, and quizzes: I do **not** accept things late. Note that you could, in theory, blow off two of these assignments and still have a perfect grade on these assignments.

This will be a course where writing down a bunch of math will **not** be sufficient to get you an A grade. Explaining, in words, what you are doing and what it means is more important. The emphasis will be on understanding, explaining, and applying, not on regurgitating a bunch of math and symbols. So, your ability to write is important.

Review questions:

I hope to provide you with review questions for each section of the course, maybe

even some answers. Knowledge of these review questions will be will be helpful when taking exams. I strongly encourage you to write out answers to all the review questions and discuss them with you classmates. You will want to form study groups. Bouncing ideas off each other will help you to determine if your thinking is correct, and will make you a better explainer.

In class I will ask many questions. I also expect you to ask questions. In addition to these question, I will sometimes give you the opportunity to earn, or lose, points by verbally answering specific questions. Participation in this latter activity is completely voluntary.

Group Assignments:

Some of the assignments will be done in groups. I will tell you in advance if an assignment is a group endeavor. The group will work together and just turn in one assignment. Everyone in the group will get the same grade for that assignment. Group activities are one of my ways of giving you an incentive to work and study together.

Office hours, contacting me outside of class, and answering questions outside of class:

My office hours will on Tuesdays from 1:00 to 2:30, and by appointment. None of you should have a class conflict with my office hours. If you can't make it to an office hour, and want to, catch me before or after class to schedule a time. My office is Econ 122. Please feel free to contact me by E-mail at Edward.Morey@Colorado.edu about setting up an appointment. Sometimes it will take a day or so for me to get back to you.

You can also E-mail me questions. I will try to answer them but will not always have the time or the knowledge to do so. If I do respond, I will typically send your question with my response to all members of the class. Information is a public good, so passing it along to everyone is, in general, efficiency increasing.

Keep in mind that if you have a question, you can E-mail your question to everyone in the class, including me. I suspect that some of your fellow students will have better answers than mine. People can comment on and add to the answers of others.