Statistics 553
Asymptotic Tools
Fall 2006

Syllabus

Syllabus - Lecture Notes - Assignments

Time and Place MWF 8:00-8:50 in 219 Thomas

Instructor David Hunter
310 Thomas
863-0979
dhunter@stat.psu.edu
Office hours: By arrangement (or just stop by).

Purpose This course will introduce students to some of the important statistical ideas of large-sample theory without requiring any mathematics beyond calculus and linear algebra. In particular, no measure theory is required. However, a basic understanding of statistics at the level of Statistics 513--514 will be assumed. Furthermore, the level of mathematical *rigor* will be high even if the level of the mathematics is not. In particular, understanding and writing proofs will be vital.

Intended Audience This course is **required** for all second-year PhD students in statistics. If you think you might be interested in taking it but you're not sure, please don't hesitate to come and talk to me or send me an email. I welcome anyone interested in this subject.

Lecture notes Lectures will be based primarily on a set of lecture notes that I will print out and bring to class. They will also be posted on the web. I will be revising these notes throughout the semester. If you have seen versions of these notes from previous semesters, be advised that many changes will have taken place. Therefore, while the previous versions will be helpful reading material, you should obtain the current versions in order to follow what is going on in class.

Required Textbook T. S. Ferguson, *A Course in Large Sample Theory* (Chapman and Hall, 1996)
This is a great book that I used to list as an optional textbook, but now I am listing it as the required textbook. I will admit that "required" is probably too strong a word, however; if you wish, you can definitely get through this class without buying the textbook because everyone will have access to the course notes (and there is a copy of the book on reserve at the PAMS library). Furthermore, we will not cover very much material from this book, though I will refer to it from time to time. However, I highly recommend the book, particularly for those of you who are pursing a PhD in statistics. First of all, it's paperback so it's not as expensive as most hardcover statistics textbooks. Second, it has the very unusual yet enormously helpful feature that the exercises are all fully worked in the appendix (the best way to learn this stuff is to work problems, and with the solutions available to guide you when you get stuck, this book is ideal for self-study). Third, it is very concisely written, managing to pack a lot more information into the average page than the Lehmann book (partly this is because it is written almost entirely in the multivariate setting, so there is no separate treatment of the multivariate case). Fourth, it is divided into small, self-contained chunks, making it possible to sample different topics in almost any order you wish. All of these features make it a terrific reference book to have on your shelf. You may wonder why, if it's such a great book,

I don't use it as the main textbook for this course. The reason is that its mathematics is a bit more advanced than I'd like, since the whole point of this course is to present as much statistics as possible without relying on too deep a mathematical background.

Optional Textbook

E. L. Lehmann, *Elements of Large-Sample Theory* (Springer, 1999)
I really like this book, especially for the level of this class. It is statistically rigorous without being overly mathematical and it contains many enlightening examples and exercises. It appears that this book is out of stock and may be hard to obtain. This won't be a serious problem, since the course notes will be available regardless. Also, there is a copy of the book on reserve in the PAMS library. I've compiled a lengthy list of errata for this book.

Computing

Computing will play a large role in the homework assignments. The software I'd recommend using is R or Splus, although I won't require any particular package or language. You can probably get by with Minitab if you're very comfortable with it, and packages such as Matlab, Maple, or Mathematica or languages such as C or Fortran should be okay as well--however, before deciding to use one of these last 5, be sure you can obtain functions like the standard normal cdf and inverse cdf as well as random deviates from not just the uniform but all the common distributions as well. You will also need to be able to produce graphics such as histograms and plots of functions. If you're not currently familiar with R or Splus, I **strongly** encourage you to visit the R project web site at www.r-project.org. There, you can download R (for free!) and obtain documentation that will teach you the rudiments of both R and Splus (for the purposes of this class, R and Splus may be considered to be the same software package; R is a free version of Splus). Go to documentation, and download "An introduction to R," which will get you started if you skim it. Start with Chapter 2 if you want a very quick introduction to the language.

Grading

There will be two midterms (15% each), a comprehensive final exam (20%), and weekly homework (50%). Most likely, the exams will be closed-book but you'll be allowed to bring a page or two of notes. This arrangement is similar to the rules for the qualifying review exam in January, which is comprised partly of questions on asymptotics. However, none of this is set in stone, so we can discuss other options if you prefer.

Integrity

All Penn State and Eberly College of Science policies regarding academic integrity apply to this course. See http://www.science.psu.edu/academic/Integrity/index.html for details.

---

dhunter@stat.psu.edu