STAT 695R: Asymptotic Statistics and Empirical Processes

Spring, 2010

Instructor: Guang Cheng, chengg@purdue.edu, Math 512
Time & Place: TU&TR, 10:30-11:45, Math 215
Office Hours: By Appointment
Prerequisites: STAT 519, STAT 528

Summary:
This *introductory level* course is composed of two sections: (I) Asymptotic Statistics (01/12-03/09, about 17 lectures); (II) Introduction to Empirical Processes (03/11-04/29, about 13 lectures).

Section 1: Asymptotic statistics is the study of large sample properties and approximations of statistical tests, estimators and procedures. In general, the goal is to learn how well a statistical procedure will work under diverse settings when sample size is large enough. We mainly use the textbook by van der Vaart (1998). **Section 1 will cover Chapters 1-2[Introduction], 3 [Delta Method], 5 [M-/Z-Estimator], 8 [Efficiency of Estimators], 10 [Bayes Procedure], 11 [Projection], 12 [U-Statistics], 21 [Quantile and Order Statistics], 23 [Bootstrap], 24 [Nonparametric Density Estimation], 25 [Semiparametric Models] of van der Vaart (1998). If time allows, I will also cover Chapters 21 [L-Statistics], 7 [Local Asymptotic Statistics] and 16 [Likelihood Ratio Test] of van der Vaart (1998).**

Section 2: The goal of second section is to introduce students with background in mathematical statistics the modern empirical processes theories. Over the past decade, the developments in empirical process theory have proven to be powerful in working with the flexible models consisting of both parametric and nonparametric components, e.g., the Cox proportional hazards model or generalized additive models. However, this increased flexibility makes distribution theory quite challenging and modern empirical process techniques are usually required. This is currently an extremely active as well as demanding area of statistical research. We will mainly use the textbook by van de Geer (2000). **Section 2 will cover Chapter 18 [Stochastic Convergence in Metric Spaces] in van der Vaart (1998), Chapters 1-2 [Introduction], 3 [Uniform Law of Large Number], 4 [Consistency], 5 [Increments of Empirical Processes], 6 [Central Limit Theorems] of van de Geer (2000). If time allows, I will also cover Chapters 7 [Convergence Rate for MLE] and 9 [Convergence Rate for LSE] of van de Geer (2000).**

Required Text:

1, van der Vaart (1998), *Asymptotic Statistics*, Cambridge U. Press

2, van de Geer (2000), *Empirical Processes in M-estimation*, Cambridge U. Press.

3, Kosorok, M.R (2008), *Introduction to Empirical Processes and Semiparametric Inference*, Springer, New York.

Recommended Text:

1, Jon Wellner's Lecture Notes (on empirical processes)
Available at http://www.stat.washington.edu/jaw/RESEARCH/TALKS/talks.html - you will need to scroll down the page.
Also check out http://www.stat.washington.edu/jaw/COURSES/EPWG/sp09.html

2, van de Geer Lecture Notes (2007)
Lectures on Empirical Processes (EMS series of lectures in mathematics)

3, van der Vaart and Wellner, J. A.(1996), *Weak Convergence and Empirical Processes: With Applications to Statistics*, Springer, New York.

4, Bickel, Klaassen, Ritov and Wellner (1993) (on semiparametric inferences). *Efficient and Adaptive Estimation for Semiparametric Models*, The John Hopkins Univ Press, Baltimore.

5, Pollard, D. (1990), *Empirical Processes: Theory and Applications.* NSFCBMS Regional Conference Series in Probability and Statistics 2. Institute of Mathematical Statistics and American Statistical Association, Hayward, California.

6, Hall, P. (1992), *The Bootstrap and Edgeworth Expansion*, Springer, New York.

7, Chen, Xiaohong (2007), *Large Sample Sieve Estimation of Semi-Nonparametric Models*, Handbook of Econometrics (edited by Heckman and Leaner), Ed 1, Vol 6, Chapter 76.

8, Dasgupta, Anirban (2008), *Asymptotic Theory of Statistics and Probability*, Springer, NY.

9, Donglin Zeng's lecture notes (on large sample theory)
See www.bios.unc.edu/~dzeng/Bios760.html

10, Barndorff-Nielsen, O. E. and Cox, D. R., (1994), *Inference and Asymptotics*, Chapman and Hall, London. (only about parametric asymptotics, not involved empirical processes)

Assignments:

No Exams

Homework:

There will be **three** homework assignments for section I and **two** homework assignments for section II. You must do all homework problems on your own, although you may discuss the problems with other students. Copying assignments or answer keys is not acceptable and is considered a violation of academic integrity. Assignments must be turned in on time for credit.

Mini-project 1:

For this project, choose a recent statistical article (published since 2000) which utilizes large sample statistical theories or empirical processes theories. *If you cannot find one, please contact me and I will assign a suitable one to you.* The major statistical or probability journals should be used. Please try to find a paper that involves a topic which seems to have significant future research potential, **write a 2-3 page summary of the paper** (you do not have to verify any math), paying particular attention to the practical issues being addressed. **The paper(s) need to be approved on or before 5pm on March 2nd. Due time for this project is 5pm on March 23rd.**

Min-project 2:

In the same paper of mini-project 1, please (i) identify 1-3 key steps in the proofs of the results which require large sample theories or empirical process theories and then verify those steps (I want you to show me that you understand the steps involved); and (ii) identify 1-3 promising problems and/or research questions which could be of interest to the statistical community. **Write a 3-4 pages summary of your verification of these steps in (i) and your findings in (ii). Also include an evaluation of the potential impact if the proposed research was successful.** You may want to get an early start on this paper(s) and meet with me several times to make sure you are comfortable with the main technical aspects of the selected paper. **Due time for this project is 5pm on May 6th.**

Final Score=

50%*(the highest three scores over 5 hwk assignments) + 20%*MP1 +

30%* MP2.