Course description: Topics will include a review of asymptotic theory for common parametric estimators and then move to theory for stochastic gradient descent as represented by 3 papers: (1) in the continuous, convex setting with iid data, (2) in the continuous, convex setting with time series data and in the non-continuous, convex setting with iid data.

Class Time and Location: 5:30-7:00pm, T,TH 1096 East Hall. No class on 10/10, 10/24, 11/11, 11/21. Makeup classes will be in Dennison 271 Fridays 5:30-7pm on 9/20, 9/27, 10/4, 11/22. Note the Fall Study Day is 10/15.

Instructor: Susan Murphy, 445D West Hall, samurphy@umich.edu


Office Hours: By Appt. on Thursday after our class.(Send me email!)


Grading: The grade is based on assigned homework and pop quizzes.

The homework is graded as follows: A random one-half of all homework problems will be graded. Ungraded homework problems will receive 2pts for being turned in. Graded homework problem will count up to 4 pts. Any extra credit homework problems will count 2pts each.

Pop quizzes will occur at the beginning of approximately 1/2 of all class periods. Each pop quiz consists of one question and counts for 1pt. Each pop quiz lasts 2 minutes. The quiz will ask a general question about the material that we are going to cover that day. For example, on the day we start the first paper below, the quiz may be to provide a one sentence definition of stochastic approximation. One way to do well on these quizzes is to read ahead and look for unfamiliar terms. Most of the time you can find definitions of these terms on wikipedia.
Advanced Topics in Theoretical Statistics
Susan Murphy (Fall, 2013)

0. A Little Review (5 lectures)
   2.1 Modes of Convergence; Skorohod representation theorem; the continuous mapping theorem; Prohorov’s Theorem (Section 2.1 in van der Vaart(1998))
   2.2 Big Oh and little Oh; Lindeberg-Feller CLT. (Sections 2.2, 2.8 in van der Vaart(1998))
   2.3 The delta method. (Chapter 3 in van der Vaart(1998))

1. Large Sample Theory for M, and Z estimators (4 lectures)
   2.1 M-estimators (like Maximum Likelihood), Z-estimators (like Estimating Equations). (Section 5.1 in van der Vaart(1998))
   2.2 Consistency of M-estimators. (Section 5.2 in van der Vaart(1998))
   2.3 Asymptotic normality of M-estimators and Z-estimators. (Section 5.3 in van der Vaart(1998))

2. Stochastic Approximation & Mirror Descent Stochastic Approximation (6 lectures).
   We will strive to cover Sections 1-3.

3. Ergodic Mirror Descent (7 lectures)
   We will strive to get up to and including Theorem 3.3.

4. Averaging Schemes for Stochastic Gradient Descent in Non-smooth Optimization (5 lectures)
   We will strive to cover Theorem 1 and Theorem 4.