

Stat 680 Syllabus – Fall 2014

Instructor	Ranjan Maitra
Office	Snedecor 2220
Phone	(515)294-7757
Email	maitra
Office Hrs	TR 1:00-1:50 pm (tentative), or by appointment
Lecture	TR 9:30-10:45
Location	Snedecor 2113
Webpage	http://www.public.iastate.edu/~maitra/stat680/fall2014.html
Textbook	None required.
Grades	Homeworks: 50%, Project: 50%.

Catalog Description

Stat 680. Advanced Statistical Computing. Cr. 3. *Prereq: STAT 543 or equivalent, and STAT 580 or equivalent.* Normal approximations to likelihoods. The delta-method and propagation of errors. Topics in the use of the E-M algorithm including; its use in the exponential family, computation of standard errors, acceleration. Resampling methods: brief theory and application of the jackknife and the bootstrap. Randomization tests. Stochastic simulation: Markov Chain, Monte Carlo, Gibbs' sampling, Hastings-Metropolis algorithms, critical slowing-down and remedies, auxiliary variables, simulated tempering, reversible-jump MCMC and multi-grid methods.

Course Information

Course Description: This course is designed for Ph. D.-level students. It is designed to survey several computer-intensive methods in statistical inference. Because of the wide range of topics to be covered, there is no prescribed text-book for the class. While the course is geared towards statistical computation, we will also look into the theory behind these methods. The broad outline for this course is as follows:

- *Review of Simulation:* Brief review of methods for sampling from given distributions: direct methods, rejection sampling, adaptive rejection. Variance-reduction methods in simulation, importance sampling and use of antithetic variates.
- *The Expectation-Maximization Algorithm, Extensions and Variants:* Theory behind Expectation-Maximization algorithm, EM in exponential families. Variance estimation for E-M and its use in Inference. Methods for speeding up EM. Alternating Expectation Conditional Maximization. Alternating Partial Expectation Conditional Maximization methods for Estimation in Mixture Models. Parallel Implementations using the Expectation Gathering Maximization. Applications.
- *Randomization tests and the Bootstrap:* Idea and techniques of Randomization tests; permutation tests, reference distributions, re-sampling methods such as jackknife and bootstrap. Parametric and non-parametric bootstrap. Estimation, confidence intervals and hypothesis testing.
- *Markov Chain Monte Carlo:* Markov Chains and Markov Random Fields; use of Markov Random Field priors in Bayesian Inference and Markov Chain Monte Carlo Methods; Gibbs sampling, burn-in, exact simulation; Multi-grid methods. Applications to Spatial Models and Image Analysis.
- *Use of High Performance Computing in Statistics:* Introduction to Parallel Programming using the HPC. Implementation of statistical methods using Message Passing Interface (MPI) and Graphical Processing Unit (GPU) computing.

Textbook: There is no required textbook for this class. However, students may find the following books useful for understanding course concepts:

- *Stochastic Simulation* by Brian D. Ripley, John Wiley and Sons, 1987.

- *Computational Statistics* by Geof H. Givens and Jennifer A. Hoeting, John Wiley and Sons, 2005.
- *The EM Algorithm and Extensions*, 2nd edition, by Geoff J. McLachlan and T. Krishnan, Wiley Interscience, 2008.
- *Numerical Analysis for Statisticians (Statistics and Computing)*, 2nd edition, by K. Lange, Springer-Verlag, 2010.
- *Parallel Programming with MPI* by Peter Pacheco, Morgan Kaufmann, 1996.

Homework: Individual practice (homework) is an important part of learning and *essential* for learning how to program. Homework assignments will be due by midnight on the due date. Solutions will be posted on the course website.

Project: There will be one very substantive final project assigned to each person during the semester. This project will involve either doing some exploratory work on a research problem, including detailed background literature study and analysis, etc or the statistical analysis of some appropriate dataset that you are interested in. You are welcome to provide this dataset in consultation with me. Another possibility would be some methodological investigation into some aspect of the topics covered in the class, or some investigation into, for instance, an engineering design problem, where the methods covered may be used to answer a question quantitatively. The final project will culminate in a written report on the level of that for a professional publication or a professional meeting in statistics and its applied disciplines. For advanced Ph. D. students, an appropriately chosen project may serve as a portion of your dissertation. Each report will be graded on the validity of the statistical analysis, the scientific component, and the quality of the write-up and presentation in communicating the results to an intended professional audience.. You are required to electronically provide all written code, documentation and datasets used in the project. Your rights, if any, to the data and software, will be preserved. *The final output should be a six-page single-spaced manuscript in 12pt font. Important dates and other format of the class project will be decided in consultation with all students.*

Computer Programming: The low-level programming language used in this class will be C. It is assumed that you are proficient in C.

Computing Framework: This class will generally use the Linux computing facilities in the Department of Statistics at Iowa State University. Additionally, we will be using Iowa State University's HPC for parallel programming. Access to the HPC is set up by policy to be dynamically controlled by Iowa State University's Information Technology Services through the class list available with the Registrar's office. Some HPC clusters to be used in the class will only be accessible by two-step authentication which includes setting up pf the Google Authenticator app to obtain the per-use verification code. Further information will be provided when the HPC account is set up.

Computer Operating System: The operating system used in this class will be Linux. The Open MPI library (<http://www.open-mpi.org>) will be used.

Course Policies

Academic Dishonesty: The class will follow Iowa State University's policy on academic dishonesty. Anyone suspected of academic dishonesty will be reported to the Dean of Students Office. <http://www.dso.iastate.edu/ja/academic/misconduct.html>.

Disability Accommodation: Iowa State University complies with the Americans with Disabilities Act and Sect 504 of the Rehabilitation Act. If you have a disability and anticipate needing accommodations in this course, please contact me (Ranjan Maitra) to set up a meeting within the first two weeks of the semester or as soon as you become aware of your need. Before meeting with (instructor name), you will need to obtain a SAAR form with recommendations for accommodations from the Disability Resources Office, located in Room

1076 on the main floor of the Student Services Building. Their telephone number is 515-294-7220 or email disabilityresources@iastate.edu . Retroactive requests for accommodations will not be honored.

Dead Week: This class follows the Iowa State University Dead Week guidelines as outlined in <http://catalog.iastate.edu/academiclife/#deadweek>.

Harassment and Discrimination: Iowa State University strives to maintain our campus as a place of work and study for faculty, staff, and students that is free of all forms of prohibited discrimination and harassment based upon race, ethnicity, sex (including sexual assault), pregnancy, color, religion, national origin, physical or mental disability, age, marital status, sexual orientation, gender identity, genetic information, or status as a U.S. veteran. Any student who has concerns about such behavior should contact his/her instructor, Student Assistance at 515-294-1020 or email dso-sas@iastate.edu, or the Office of Equal Opportunity and Compliance at 515-294-7612.

Religious Accommodation: If an academic or work requirement conflicts with your religious practices and/or observances, you may request reasonable accommodations. Your request must be in writing, and your instructor or supervisor will review the request. You or your instructor may also seek assistance from the Dean of Students Office or the Office of Equal Opportunity and Compliance.

Contact Information: If you are experiencing, or have experienced, a problem with any of the above issues, email academicissues@iastate.edu.

Grading: Letter grades including plus/minus will be given based on performance on homework and projects. The specific grading scale is not determined until after all grades have been calculated.

Course Prerequisite: STAT 543 or equivalent and STAT 580 or equivalent. There are no exceptions. A student not having all the prerequisites will receive an automatic F.

August 28, 2014