STAT 606: Computational Statistics (Winter 2008)

- **Web site**: ctools.umich.edu.
- **Instructor**: Yves Atchadé (yvesa@umich.edu); Office: 445D West Hall, Phone: 763-5238.
- **Office Hours**: By appointment or just stop by.
- **Lectures**: Tu-Th. 1-2:30pm, B760 Est Hall. First lecture Jan. 4, last lecture April 15.
- **Prerequisites**: Regression methods at the level of Stat 500; probability and statistical theory at the level of Biostat 501/602. Some exposure to computer programming is strongly recommended.
- **Objectives**: This course covers the use of computer algorithms in statistical modeling. The emphasis is on both the understanding and the practical implementation of the algorithms. At the end of the course, the student should be able to understand the building principles and the practical aspects of most algorithms routinely used in statistical research.
- **Computing**: The class will be taught using programs written in Matlab and R. Students may write their homework solutions using R, Matlab or C. SPSS, SAS and other very high level statistical packages are not suitable for this course.
- **Outline**:

  **Part I**  Numerical linear algebra: Basic linear algebra; SVD; QR Factorization; Application to the Least Square Method.

  **Part II**  Basic Monte Carlo Methods: Random number generation, inversion method, rejection method, importance sampling, variance reduction.

  **Part III**  Markov Chain Monte Carlo methods: Markov chains; Metropolis-Hastings algorithm, Gibbs sampler; Data augmentation, EM algorithm; Sequential Monte Carlo methods; Application to Bayesian inference.

  **Part IV**  Optimization methods: line-search, Newton methods, Conjugate gradient, trust-region; Applications to Maximum likelihood inference; Constrained optimization.

- **Grading**: Homework (50%), Project: (50%). The homeworks will be a mix of theoretical questions and computer programming. There will be no exam.

- **Final Project**: The final project for this course will require each student to review a research article related to this course and write a report. A list of possible articles will be provided. Students can, and are encouraged to propose other relevant papers; may be in connection with their research interests. This project will have two parts.

  **Part 1**: By March 06, 2008, each student should have decided on at least two papers. For each selected paper, the student will write a very brief description of the problem and a proposal for a simulation study. The total length of this mid-term report should be about 1 page for each article and should not exceed 2 pages for each article.

  **Part 2** By the end of the semester (April 15), each student will review his/her chosen paper.
(among the two initial selections) and submit a typed report. The length of the report should not exceed 15 pages including all tables and figures. Each report should consist of the following:

- An introduction to the problem considered in the paper.
- A description of the approach taken by the author(s).
- The relation of this paper to the literature.
- A description of the simulation study to be performed by the student. For example, the student might choose to verify the simulation results of the paper, examine an unexplored special case or extension, or apply the method of the paper to a similar problem. Other ideas are welcome.
- The results of the simulation study.
- A discussion and a conclusion.
- The references.