STAT 7650: COMPUTATIONAL STATISTICS
Course Syllabus

CONTACT INFORMATION:
Instructor Elvan Ceyhan
Office 203 Extension Hall
Office Hours Tu 11:00 am - 12:00 pm or by appointment
Phone (334) 844 3613
E-mail ceyhan@auburn.edu (best way to contact me)

TIME & PLACE: Tu-Th 9:30 am - 10:45 am SCC (Science Center Classroom Bldg.) 122

Description and Objectives:
Description:
Computational Statistics focuses on numerical methods for addressing statistical problems. This course introduces modern computational techniques, including numerical optimization (e.g., EM algorithm, Fisher scoring), numerical integration (e.g., quadrature, Monte Carlo), and Bayesian inference methods (e.g., MCMC, variational inference). Applications span statistical machine learning, computational biology, and other fields. Coursework includes computational assignments.
Objectives:
By the end of the course, students will:
1. Apply numerical optimization techniques to solve complex statistical problems.
2. Analyze the strengths and limitations of numerical integration and Bayesian inference methods in real-world applications.
3. Evaluate the scalability and efficiency of advanced computational methods for handling large datasets.
4. Create innovative solutions using integrated computational techniques.
5. Synthesize knowledge to develop scalable Bayesian methods and enhance statistical model performance.
6. Critique the impact of computational statistics on scientific discovery.

TEXT: Required:
Givens, G. H. and Hoeting, J. A. (2012) *Computational Statistics*, 2nd Edition, Wiley-Interscience. Students are strongly recommended to purchase a copy of this book.

Recommended:
- Gentle, *Computational Statistics*, Springer, 2009
- Efron and Hastie, *Computer Age Statistical Inference*, Cambridge, 2021

PREREQUISITES: STAT 7020 (Regression Analysis) and STAT 7610 (Statistical Theory and Methods II). Some background in probability and prior experience programming in R is helpful. The level of mathematics needed in this course does not extend much beyond the Taylor series and linear algebra. Breadth of mathematical training is more helpful than depth. The expected level of statistics is equivalent to that obtained by a graduate student in his or her first year of study of the theory of statistics and probability. An understanding of maximum likelihood methods, Bayesian methods, elementary asymptotic theory, Markov chains, and linear models is most important. Refresher material on the statistical methods will be provided on Canvas in the first week of the semester. Also, Chapter 1 of the textbook reviews notation and provides background material in mathematics, probability, and statistics.

COURSE WEBSITE: posted under *Canvas*

SOFTWARE: The R software will extensively be used in this class. R can be downloaded for

free from http://www.r-project.org or RStudio from https://www.rstudio.com. *Use other software at your own risk.*

HOMEWORK: At least one homework will be assigned at the conclusion of each chapter. Due to the computational nature of HW in this course, it is highly encouraged to prepare HW assignments in R markdown, or some similar software. Homework could be completed in groups of 2-3 students each (but you must acknowledge at the beginning of the submitted work).

EXAMS (Tentative): There will be one midterm and one final exam. MT is scheduled to be in the middle of the Spring Semester. If MT is during the class time in the regular classroom, one 8.5×11 formula sheet (both sides) is allowed for each midterm (MT) exam and two such sheets (both sides for each) are allowed for the final exam.

GRADING: Homework (40%), MT (30%), Final (30%). The grading scheme is [85-100] = A, [70-85) = B, [55-70) = C, [40-55) = D, and [0-40) = F.

COURSE CONTENT (Tentative):

Optimization Methods: 1. Optimizing smooth univariate functions: Bisection, Newton's method, Fisher scoring, secant method, scaled fixed point iteration 2. Optimizing smooth multivariate functions: Newton's method, Fisher scoring, Gauss-Newton method, ascent algorithms, discrete Newton method, scaled fixed point iteration, quasi-Newton methods, nonlinear Gauss-Seidel iteration, Nelder-Mead algorithm *3. Combinatorial optimization: Local search, ascent algorithms, random starts local search, tabu algorithms, simulated annealing, genetic algorithms 4. EM Optimization strategies: EM, SEM, MCEM, ECM, EM gradient, accelerated EM strategies.*

Numerical and Monte Carlo Integration Methods: 5. Numerical integration: Newton–Cotes and Gaussian quadrature, frequently encountered problems 6. Simulation & Monte Carlo Integration: Monte Carlo Integration; variance reduction techniques: importance sampling, antithetic sampling, control variates, Rao-Blackwellization; sampling importance resampling (SIR) 7. Markov Chain Monte Carlo Methods: Metropolis-Hastings algorithm, Gibbs sampling, implementation issues 8. Advanced MCMC Methods: Adaptive MCMC, Reversible jump MCMC, Slice sampling and other auxiliary variable methods, Perfect sampling

Other Topics (time permitting)

Bootstrap; Nonparametric smoothing; More advanced MCMC, Langevin Metropolis-Hastings Algorithm, Hit-and-run algorithm, Multiple-try Metropolis-Hastings Algorithm, and more; Alternatives to MCMC; Variational Bayes, Integrated Laplace Approximation (INLA); Approximate Bayesian Computation (ABC).

Course Policies

Attendance: You are strongly encouraged to attend classes on a regular basis. You will be held responsible for all the material that is covered in class.
Working Together: You may work together on homework, and are encouraged to do so, unless otherwise instructed. You must, however, write up and submit your own work.
Makeup: You will not be allowed to make-up missed exams except under extraordinary circumstances. If you know that you will miss an exam, please let me know in advance.
Late Work: Late homework/lab has a half-life of 24 hours; that is, you receive 50% credit if work is submitted within 24 hours of the due time. You will not receive credit after that.

- **Policies on Class Attendance, Submission of Late Written Assignments, Missed In-Class Work and Missed Examinations:**
  - **Excused Absences:** Students are granted excused absences from class for the following reasons: Illness of the student or serious illness of a member of the student's immediate family, death of a member of the student's immediate family, trips for student organizations sponsored by an

academic unit, trips for University classes, trips for participation in intercollegiate athletic events, subpoena for a court appearance and religious holidays. Students who wish to have an excused absence from this class for any other reason must contact the instructor in advance of the absence to request permission. The instructor will weigh the merits of the request and render a decision. When feasible, the student must notify the instructor prior to the occurrence of any excused absences, but in no case shall such notification occur more than one week after the absence. Appropriate documentation for all excused absences is required.

- **Make-Up Policy:** We jointly will decide to take the proper action and/or make up missed major examinations (e.g., hour exams, mid-term exams) due to properly authorized excused absences. Except in unusual circumstances, such as continued absence of the student or the advent of University holidays, a make-up exam will take place within two weeks from the time the student initiates arrangements for it. Except in extraordinary circumstances, no make-up exams will be arranged during the last three days before the final exam period begins. The format of the make-up exam will be (as specific by the instructor).

- **Email and Canvas Use:** Students have control of the notification settings in their Canvas accounts. You should set up their notifications to alert them when an Announcement is posted, an Assignment is due, a grade is released, etc. For students new to Canvas, save time (and emails) by sharing this link to a 7 minute "Getting Started with Canvas" video (and transcript) (Links to an external site.) created by Canvas LMS. In order to protect your privacy, all course e-mail correspondence must be done through a valid AU account.

- **Provisions of the Americans with Disabilities Act:** Students who need accommodations are asked to electronically submit their approved accommodations through AU Access and to make an individual appointment with the instructor during the first week of classes – or as soon as possible if accommodations are needed immediately. If you have a conflict with my office hours, an alternate time can be arranged. To set up this meeting, please contact me by e- mail. If you have not established accommodation through the Office of Accessibility, but need accommodation, make an appointment with the Office of Accessibility, 1228 Haley Center, 844-2096.

- **Classroom Behavior:** The Auburn University Classroom Behavior Policy is strictly followed in the course; please refer to the Student Policy eHandbook for details of this policy and the Policy on Classroom Behavior.

- **Emergency Contingency Plan:** If normal class and/or lab activities are disrupted due to illness, emergency, or crisis (such as an H1N1 flu outbreak), the syllabus and other course plans and assignments may be modified to allow completion of the course. If this occurs, an addendum to your syllabus and/or course assignments will replace the original materials.

- **Academic Dishonesty:** Auburn University expects that all students have read and understood the University's Code of Student Conduct and that all students will complete all assignments with fairness and honesty. Failure to follow the rules in the University's Code of Conduct may result in academic misconduct and being reported to the appropriate committee. If the committee determines an academic misconduct happened, the sanctions for the act can include a failing grade in the course and suspension or dismissal from the University.

- **AI Usage Policy**

  AI tools (e.g., ChatGPT) will be permitted in this class to support your work, with some assignments requiring their use.

  To maintain academic integrity, students must disclose any AI-generated material they use and properly attribute it, including in-text citations, quotations, and references (such attribution and acknowledgement is also required any other online tool or book or written work). Students should exercise caution and avoid sharing any sensitive or private information when using these tools. Examples of such information include personally identifiable information (PII), protected health

information (PHI), financial data, intellectual property (IP), and any other data that might be legally protected.
A student should include the following statement in assignments to indicate use of a Generative AI Tool: “The author(s) would like to acknowledge the use of [Generative AI Tool Name], a language model developed by [Generative AI Tool Provider], in the preparation of this assignment. The [Generative AI Tool Name] was used in the following way(s) in this assignment [e.g., brainstorming, grammatical correction, citation, which portion of the assignment].”

Summary of Guidelines for AI Use:
• Effort matters: Refine prompts to achieve meaningful results.
• Verify accuracy: Validate any numbers or facts provided by AI with trusted sources.
• Acknowledge use: Include a paragraph in assignments detailing how AI was used and the prompts employed. Failure to do so violates academic honesty policies.
• Use thoughtfully: Employ AI only when appropriate for the task.
This policy is aimed to convey the idea that AI is used as an effective tool while promoting critical thinking and accountability.