UCLA Department of Statistics

Stats 210 Computer Intensive Methods

Instructor: Qing Zhou [zhou@stat.ucla.edu], OH: TR 4:45-5:30pm, MS 8979.
Recommended Requisites: Stats 200B, programming skills (R, C/C++, Matlab, etc.).

Overview

This is an elective course for graduate students in statistics and related fields. Computer-based methods of statistical inference can answer many real-world statistical questions without complicated mathematics. This course will be a good preparation for graduate students interested in data science research.

Course Objectives

The goal of this course is to arm students with computational techniques that they can use to analyze and understand complicated data sets. At the completion of this course, the student will be able to:

- Apply bootstrap methods to make statistical inference via computer simulation.
- Develop resampling methods to quantify the uncertainty in statistical estimation.
- Implement Monte Carlo algorithms to simulate from probability distributions and calculate expectations.
- Develop efficient algorithms to evaluate the accuracy of statistical methods for complicated real-world data.
- Perform simulation-based inference on large data sets via data splitting and resampling techniques.

Course Description

An overview of the theory and practice of computer-based methods for statistical inference and uncertainty quantification, including the bootstrap, resampling, computer simulation and Monte Carlo sampling. Coverage of nonparametric and parametric bootstrap, bootstrap inference, permutation test, cross-validation, likelihood approximation, importance sampling, Markov chain Monte Carlo and applications of these methods in statistical inference and scientific computing.

Topics & Tentative Schedule

Below is a tentative schedule of the course with weekly topics:

- Week 1: Introduction and review of statistical inference. Bootstrap standard errors.
- Week 2: More complicated data structures. Regression models.
- Week 3: Estimates of bias via bootstrap.
- Week 4: Confidence intervals. Interval estimation by the bootstrap.

- Week 5: Permutation tests. Testing with the bootstrap.
- Week 6: Introduction to Monte Carlo methods: Inverse-cdf method, rejection sampling, and exact simulation for chain structures.
- Week 7: Importance sampling, sequential Monte Carlo, particle filter.
- Week 8: Markov chain Monte Carlo: The Metropolis-Hastings algorithm and the Gibbs sampler, Bayesian Missing data problem.
- Weeks 9-10: Student oral presentations of selected topics.

Assignments

There will be four homework assignments which include both theoretical problems and applied problems that require computer implementation and data analysis.

Grading

Final grade of this course is composed of threes parts:

1. Homework assignments (60%).
2. Final presentation (40%): Groups of no more than four students. A topic from the reference books (see below) will be assigned to each group. Each group will prepare and submit the slides of their talk, and every group member must present a portion of the slides.

References

- Lecture notes: To be posted on the CCLE site weekly.
- Efron, B. and Tibshirani, R.J., An Introduction to the Bootstrap (1994), Chapman & Hall.
- Liu, J.S., Monte Carlo Strategies in Scientific Computing (2008), Springer.