

统计模型与计算方法

Time and Location

每周一 7 8 节，双周三 3 4 节，三教 501

Course description

The course develops computational methods for statistical estimation and inference: random-variable generation, Monte Carlo integration, importance sampling, likelihood optimization, EM, data augmentation, MCMC, Kalman and particle filtering, resampling, robust statistics and regression, causal graphical algorithms, and causal effect estimation. Emphasis is placed on the statistical target, algorithmic derivation, assumptions, implementation in R, and interpretation of the results.

Prerequisites and learning outcomes

Multivariable calculus, linear algebra, probability, mathematical statistics, and basic R programming. Students should review conditional expectation, LLN/CLT, likelihood, and linear regression. Specialized prior courses in Bayesian computation or causal inference are not required.

By the end of the course, students should be able to:

1. Derive representative simulation, optimization, MCMC, filtering, and resampling algorithms.
2. Implement statistical algorithms and diagnose failure or inappropriate application.
3. Formulate and compute robust estimators and explain their limitations.
4. Query causal DAGs, find adjustment sets, interpret structure-learning output, and estimate causal effects under stated assumptions.

Texts and their roles

- **GH (primary):** Givens, G. H. and Hoeting, J. A. *Computational Statistics*, 2nd ed., Wiley. First published 2012; copyright 2013. Monte Carlo, optimization, EM, MCMC, SMC, and bootstrap.
- **LD (Chinese companion):** 李东风, 《统计计算》, 2022-08-29 讲义版。R examples, generation, Monte Carlo, MCMC, SMC, and resampling.
- **RZ:** Rizzo, M. L. (2008). *Statistical Computing with R*, 1st ed., Chapman & Hall/CRC. Generation, Monte Carlo methods, jackknife, bootstrap, and permutation tests.
- **TA:** Tanner, M. A. (1993). *Tools for Statistical Inference*, 2nd ed., Springer. EM, data augmentation, Gibbs sampling, and MCEM extensions.
- **HB:** Gentle, J. E., Härdle, W. K., and Mori, Y., eds. (2012). *Handbook of Computational Statistics: Concepts and Methods*, 2nd revised and updated ed., Springer. Selected regression material and the Davies–Gather chapter on robust statistics.

Workload and assessment

Four problem sets: 60% (15% each). Final project: 40%。

Problem sets combine mathematical derivation, implementation, diagnostics, and computational experiments. Discussion of general ideas is allowed, but submitted work must be the student's own; external sources, software, and AI assistance must be acknowledged.

R is the main language. Students must explain their algorithms, assumptions and results, and acknowledge external code and assistance under the confirmed course policy.