Statistical Computing and Simulation

Spring 2025

Instructor: Jack C. Yue (余清祥)
Lecture Times: 9:10~12:00 Tuesday
Phone: 2938-7695, e-mail: csyue@nccu.edu.tw (strongly recommended)
Home page: http://csyue.nccu.edu.tw
Office Hours: 14:00~16:00 Tuesday, or by appointment

References:

1. R related references
   → ggplot2: Elegant Graphics for Data Analysis (2016) by H. Wickham
   Data Visualization: A Practical Introduction (2018) by K. Healy
2. Statistical Computing
   → Elements of Statistical Computing (1988) by R.A. Thisted
   Numerical Methods of Statistics (2001) by J.F. Monahan
   Handbook of Computational Statistics: Concepts and Methods (2004) by J. E. Gentle, W. Härdle, and Y. Mori (Eds.)
3. Computer Simulation
   → Stochastic Simulation (1987) by B.D. Ripley
   A Course in Simulation (1990) by S.M. Ross
   Modern Simulation and Modeling (1998) by R.Y. Rubinstein & B. Melamed
   Simulation and the Monte Carlo Method (1981) by R.Y. Rubinstein
4. Manuals and References at www.r-project.org and at my home page

Course Description:

Mathematical analysis was used to be the most useful tool, and probably the only tool, in handling statistical problems. The rapid development of computers in recent years has made simulation a powerful tool as well, and it is especially convenient in dealing with problems without "good" statistical assumption. However, simulation is like mathematical experimentation, it needs careful design and planning in order to come out with satisfied results. At the first half of this course, we will introduce basic principles of computing and simulation, including generation of random numbers and random variables, and statistical tests. Advanced techniques and applications shall be covered in the second half of the semester. Topics covered in this course include: *Simulation and Monte Carlo methods*, *Matrix computation*, *Numerical integration and approximation*, *Data partition and resampling*, *Optimization methods*, *Density estimation*, and *Bayesian computing*. Also, the use of statistical software R/S-Plus is required in this course. The software R can be downloaded via http://www.r-project.org.

Grading:

Grades will be based on regularly assigned homework plus a project. The project will be due on the final week of spring semester (6/23/2024, Friday).

Topics and contents:

- **Simulation and Monte Carlo methods**
  →Pseudo-random number generation, Linear congruential method, Inverse method, Rejection method, and Statistical tests
- **Matrix computation**
  →Least square methods, Gram-Schmidt method, Gaussian elimination, Singular value decomposition, Cholesky decomposition
- **Numerical integration and approximation**
  →Trapezoidal and Simpson's rules, General Newton-Cotes rules, Monte-Carlo integration
- **Data partition and resampling**
  →Bias reduction, Variance estimation using Jackknife and Bootstrap (including Dependent Data and Bootstrap), MCMC (Markov Chain Monte Carlo)
- **Optimization methods**
  →Maximum likelihood estimation, Newton-Raphson and Newton like methods, Fisher scoring methods, EM algorithm
- **Density estimation**
  →Histograms and related density estimator, Spline smoothing, Kernel smoothing
- **Bayesian computing**
  →Bayes' Theorem, Bayesian thinking, Bayesian computation, Markov Chain Monte Carlo methods

Class Handouts:

Class materials, including in-class handouts, will be posted on my website http://csyue.nccu.edu.tw and no hard copies will be distributed.

- S-Plus and R users' guides and notes
- Class files (Microsoft power point or Acrobat pdf format)
- Class related papers and reports
- Homework assignments and solutions

Note: The homework is usually on a 2-week interval base and due on Tuesday/ Friday afternoon at 5. However, you need to hand-in your homework and final report in hard copy, and no email copies are allowed.