Lecture 27: Wrap-up

Pratheepa Jeganathan

06/05/2019

Final Project

- **Due June 5, 2019 by 11:59 PM**
- The project report can be 10-12 pages long (excluding **bibliography that should be very complete**).
- Zip all your files and name the compressed folder as YourFirstNameYourLastName.
- Send the zip folder to pjeganat [at] stanford [dot] edu; **Write in the subject line** STATS 205 Final Project.

Course Evaluations Now Open

- Axess is now open to complete end-term course evaluations.
- You can find it on
  - Stanford Axess
  - in Course and Section Evaluations
  - on the Student tab
- You may complete the evaluations by 06/17/2019.
- Grades are available by 06/18/2019.

Grading

- The final letter grade for this course will be determined by each method of assessment weighted as follows:
  - Class participation (5%)
  - Weekly homework assignments (50%)
  - Midterm project proposal (10%, due on 05/03/2019)
  - Final project (35%, due on 06/05/2019)

Expected outcomes

By the end of the course, the student should be able to

1) understand the assumptions underlying the nonparametric methods
   - rank-based methods for parameter estimation, confidence intervals, and hypothesis testing in one-sample, two-sample, ANOVA.
   - rank-based methods for discrete data problems.
   - nonparametric bootstrap for testing and confidence intervals, better bootstrap confidence intervals, permutation method for hypothesis testing, jackknife for estimating standard error of some estimators.
   - rank-based methods for testing association.
   - rank-based linear regression, nonparametric regression, wavelets.
   - statistical functionals, influence functions.
   - data visualization tools for data exploration in nonparametric settings (association plots, mosaic plots, median polish, Tukey additivity plot).

Expected outcomes

2) apply nonparametric methods to modern data analysis problems
   - final project.
   - homework problems.

Expected outcomes

3) get hands-on experience in implementing methods and using existing R packages
   - All lectures R Markdown files are shared with students.
   - All the in-class examples are illustrated with R packages.
   - We wrote R functions when there is no built-in functions.
     - permutation.
     - bootstrap (some examples).
     - some simulations.
   - All homework solutions were written using R and R Markdown.
   - Final project is done using R packages.

Textbook

Additional topics that were not covered in this class.

- Ranked set sampling (RSS) (**HWC** Chapter 15).
  - a technique of data collection that generally leads to more efficient statistical procedures than competitors based on simple random samples (SRS).
- Introduction to survival analysis (**HWC** Chapter 11).
  - Statistical methods for incomplete (censored) data.
- Introduction to Bayesian nonparametric (**HWC** Chapter 16).
  - In parametric Bayesian inference
    - $Y_1, Y_2, \cdots, Y_n \sim f(y|\theta)$. We put prior $\Pi(\theta)$ on parameter θ.
  - In nonparametric Bayesian inference
    - Replace the finite dimensional model $\{f(y|\theta) : \theta \in \Theta\}$ by infinite dimensional model such as
      $\mathbb{F} = \left\{ f : \int \left(f''(y) \right)^2 dy < \infty \right\}$.

Additional notes

- Statistics for research in biology (more than this)
  - Susan Holmes Lab
- Learn about applied statistics (Stats305)
  - Susan Holmes
- Bayesian statistics
  - Learn R package Rstan.
  - Statistical Rethinking by Richard McElreath.