

STAT 5010/PSYC 6120
Introduction to Nonparametrics & Log-Linear Models
Section 401

Instructor: Wei Wang wwa@upenn.edu
Class hours: Tuesday and Thursday, 12:00 pm - 1:30 pm (EST).
Office hours: After class, email or by appointment.

Course description

The course covers commonly used nonparametric (semiparametric) statistical techniques and various regression models. Topics include the Wilcoxon rank sum test, signed rank test, the Kruskal-Wallis test, two sample tests on proportions, smoothing methods (kernel smoothing and spline smoothing), generalized linear models and generalized additive models.

References

Nonparametric Statistical Methods, M. Hollander, D. A. Wolfe, and E. Chicken.
An Introduction to Categorical Data Analysis, A. Agresti.
Nonparametric Statistical Methods Using R, J. Kloeke and J. W. McKean.

Prerequisite: STAT 5000.

Undergraduate level probability and statistics: conditional probability, random variables, distributions, sample mean and sample variance, central limit theorem, point estimation, hypothesis testing, p-value, confidence intervals, and multiple linear regression.

Calculus: continuity, differentiation, and integration.

Linear algebra: vectors, matrices, and matrix multiplication.

R programming experience: scatter plots, histograms, and data summary.

Software

We will use the free statistical computing software R (<http://www.r-project.org/>) frequently in class. Sample R code will be provided to help you solve homework problems.

Homework

The homework will be assigned biweekly. Without a convincing reason, late homework will not be given full credit (25 points off every 24 hours). If you are not certain about your situation, ask the instructor in advance instead of a last minute request.

Grading

The final grade will be based ~60% on homework, and 40% on the take home final exam.

Notes

- Pdf format is preferred, and html format is not recommended.
- If R is used, attach both the code and the output. Otherwise, 50% points will be deducted. Irrelevant R output will be penalized.
- The solution must be in your own words. Identical solutions will be marked zero.
- Try to provide some explanation of your answer instead of a simple yes or no.
- You are not allowed to use generative AI (e.g, tools like ChatGPT) for your work. Using such tools in this course will be considered a violation of Penn's Code of Academic Integrity and will be reported to the Center for Community Standards and Accountability.
- Regular classroom attendance and participation is anticipated. If you are not able to attend some lectures, you can request to view the recorded videos. The request will be approved, within reason.
- Independent work is expected for the final exam. Discussion or communication with other people is not allowed. Otherwise, it will be considered cheating, and the exam paper will be marked 0.