Bayesian and Network Statistics

Syllabus

NETS 7350
Northeastern University
Th 1:00pm - 4:00pm
177 Huntington, 2nd Floor

Professor: Nick Beauchamp
Email: n.beauchamp@northeastern.edu
Office: 931 Renaissance Park
Office Hours: MW: 1:00pm-2:30pm
(Please try to make an appointment first via email to ensure I can fit you in.)

Course Description

This course offers an introduction to advanced quantitative methods including maximum likelihood, hierarchical models, sampling, and network modeling. Students with some experience in basic econometric methods will learn to estimate and develop models from the probabilistic and Bayesian perspective, and will pursue their own research project along the way, with a particular attention to the methodological challenges. The course begins with a review of probability, then examines maximum likelihood methods for estimating regression models with continuous and categorical dependent variables. This is followed by examining a variety of procedures for sampling from posterior distributions, including grid, Gibbs, and Metropolis sampling. These methods are then applied to hierarchical modeling and other simple probabilistic models. The course then takes a closer look at the statistical modeling of networks as it has been developed in the social sciences, including Exponential Random Graph Models (ERGM), temporal models such as TERGM and SIENA, spatial network models, and stochastic block models. Along the way, students will develop their own modeling projects using datasets of their own devising, which will be presented, reviewed, and discussed in detail in the latter third of the course, giving close attention to methodological issues and potential solutions.

Requirements

Students are expected to have taken a prior graduate-level course in statistics or econometrics. For this course, the main project is a final paper (worth 50% of the grade) in the student's preferred field of study, with an emphasis on crafting, implementing, and testing the appropriate statistical models. There will also be three problem sets, each worth 10% of the grade, which will focus primarily on preliminary data analysis and exploration. The remaining 20% of the grade will be presentations (10%) and participation (10%).

Required texts

1. *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists.* Scott M. Lynch, 2007.
2. *Statistical Rethinking: A Bayesian Course with Examples in R and Stan.* Richard McElreath, 2014.

Other useful texts:

1. *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan* (Second Edition). John Kruschke, 2015.
2. *Bayesian Analysis for the Social Sciences.* Simon Jackman, 2009.
3. *Bayesian Data Analysis* (Third Edition). Andrew Gelman et al, 2013.

For many other useful suggestions, see also here: http://stats.stackexchange.com/questions/125/what-is-the-best-introductory-bayesian-statistics-textbook

Schedule

Note that topics map only approximately onto weeks, depending on time and student interest. Problem sets will roughly comprise topics 1-2, 3-4, and 5-6. Students will also be asked to present at least one research article of their choosing using Bayesian methods in their area of interest.

1. Probability Review
2. Maximum Likelihood
   Lynch, Ch 2; McElreath, Ch 1.
3. Bayesian statistics.
   Lynch, Ch 3; McElreath, Ch 2.
4. Sampling 1: Grid, Inversion, Rejection, Importance.
   Lynch, Ch 4; McElreath, Ch 3.
5. Sampling 2: Metropolis-Hastings and Gibbs.
   Lynch, Ch 5-6; McElreath, Ch 8.
6. Linear regression and probit.
   Lynch, Ch 7-8; McElreath, Ch 4-5, 9-10.
7. Hierarchical and spatial models.
   Lynch, Ch 9; McElreath, Ch 12-13.
8. Spatial network models and stochastic block models

9. Exponential Random Graph Models
10. TERGM and Siena
11. Project discussions and presentations

For the paper, I will be asking for:

1. Three one-paragraph paper proposals (Around week 4)
2. A three-page research plan, including some lit review (Around week 7)
3. Short presentations of preliminary status (Around week 10)
4. Longer presentation of paper results (Last week or two)
5. Final paper submission (April 24)

Learning outcomes

Upon completing this course, students will be able to:

- Demonstrate an understanding of probability, Bayesian statistics, sampling methods, probabilistic model design, and network models.
- Synthesize, analyze, and critically evaluate major arguments in contemporary research papers using Bayesian and network statistics.
- Develop and apply Bayesian methods to their own modeling designs.
- Communicate effectively in written and oral formats relevant to Bayesian and related network methods.
- Design and complete an independent research project that employs probabilistic modeling.

Policies and Accommodations

Incompletes Except in the most serious circumstances, Incompletes in this course are not possible. If the professor agrees to an Incomplete, a form in the Political Science Department must be filled out, representing a contract between the student and the faculty member on when and how the course will be completed.

Accommodations for students with disability/ADA Northeastern is fully committed to creating a community characterized by inclusion and diversity. As part of this commitment, it upholds the American with Disabilities Act as Amended of 2008 and the American with Disabilities Act and Section 504 of Rehabilitation Act, referred to collectively as the ADA. The ADA requires Northeastern to provide reasonable accommodations to students with disabilities unless doing so would create an undue hardship, compromise the health and safety of members of the university community, or fundamentally alter the nature of the universitys employment mission. Students seeking information regarding ADA accommodations should review the Universitys ADA Information and Resources Procedure available here.

Academic Integrity The Department of Political Science takes very seriously the issue of academic honesty, and as set forth in Northeastern Universitys principles on Academic Honesty and Integrity Policy (the complete text can be found at NEUs Office of Student Conduct and Conflict Resolution). Any student who appears to violate these principles will fail the course and will be put on academic probation. Individual faculty, with the support of the Department, can impose harsher penalties and as they deem necessary. Cheating is one example of academic dishonesty, and which is defined as using or attempting to use unauthorized materials, information, or study aids in any academic exercise. When completing any academic assignment, a student shall rely on his or her own mastery of the subject. Cheating includes plagiarism, which is defined as using as ones own the words, ideas, data, code, or other original academic material of another without providing proper citation or attribution. Plagiarism can apply to any assignment, either final or drafted copies, and it can occur either accidentally or deliberately. Claiming that one has forgotten to document ideas or material taken from another source does not exempt one from plagiarizing. Your instructor will clarify specific guidelines on fair use of material for this class.

Title IX Northeastern is committed to providing equal opportunity to its students and employees, and to eliminating discrimination when it occurs. In furtherance of this commitment, the University strictly prohibits discrimination or harassment on the basis of race, color, religion, religious creed, genetic information, sex, gender identity, sexual orientation, age, national origin, ancestry, veteran, or disability status. The Northeastern University Title IX policy articulates how the University will respond to reported allegations of sexual harassment involving students, including sexual assault, and provides a consolidated statement of the rights and responsibilities under University policies and Title IX, as amended by the Violence Against Women Reauthorization Act of 2013.