Syllabus

Course goals & objectives:

This course introduces students to linear models and its extensions for model building, including exploratory data analysis techniques, variable transformations and selection, parameter estimation and interpretation, prediction, hierarchical models, model selection and Bayesian model averaging. The concepts of linear models will be covered from Bayesian and classical viewpoints. Topics in Markov chain Monte Carlo simulation will be introduced as required, however it is expected that students have either taken STA 601 or are co-registed.

All students should be comfortable with linear algebra and mathematical statistics at the level of STA 611.

The course goals are as follows:

1. Understand the different philosophical approaches to statistical analyses (Bayesian and frequentists)
2. Build a solid foundation for the probability theory of Gaussian linear models and hierarchical models.
3. Build appropriate statistical models for data perform data analysis using appropriate software, and communicate results without use of statistical jargon.

Topics

Course topics will be drawn (but subject ot change) from

- Motivation for Studying Linear Models as Foundation
- Random Vectors and Matrices
- Multivariate Normal Distribution Theory
- Conditional Normal Distribution Theory

- Linear Models via Coordinate free representations (examples)
- Maximum Likelihood Estimation & Projections
- Interval Estimation: Distribution of Quadratic Forms
- Gauss-Markov Theorem & Optimality of OLS
- Formulation of Bayesian Inference
- Subjective and Default Priors
- Related Shrinkage Methods and Penalized Likelihoods (Ridge regression, lasso, horseshoe etc)
- Model Selection (comparison of classical and Bayesian approaches)
- Bayes Factors
- Bayesian Model Averaging
- Model Checking: Residual Analysis, Added-Variable Plots, Cooks-Distance Transformations
- Bayesian Outliers
- Bayesian Robust Methods for Outliers
- Generalized Linear Model and Weighted Regression
- Hierarchical Models

Grading:

Homework	20%
Midterm	25%
TakeHome	25%
Final	25%
Participation	5%

Grades may be curved at the end of the semester. Cumulative numerical averages of 90 - 100 are guaranteed at least an A-, 80 - 89 at least a B-, and 70 - 79 at least a C-, however the exact ranges for letter grades will be determined after the final exam. The more evidence there is that the class has mastered the material, the more generous the curve will be.

Homework:

These will be assigned at each class or weekly on the course webpage.

The objective of the problem sets is to help you develop a more in-depth understanding of the material and help you prepare for exams and projects. Grading will be based on completeness as well as accuracy. In order to receive credit you must show all your work.

Lowest score will be dropped.

You are welcomed, and encouraged, to work with each other on the problems, but you must turn in your own work. If you copy someone else's work, both parties will receive a 0 for the problem set grade as well as being reported to the Office of Student Conduct (http://www.studentaffairs.duke.edu/conduct). Work submitted on Sakai will be checked for instances of plagiarism prior to being graded.

Submission instructions: You will submit your HW on Sakai by uploading a PDF. If the TAs cannot view your work, or read your handwriting, you will lose points accordingly.

All assignments will be time stamped and late work will be penalized based on this time stamp (see late work policy below).

Attendance & Participation:

You are expected to be present at class meeting and actively participate in the discussion. Your attendance and participation during class, as well as your activity on the discussion forum on Sakai will make up 5% of your grade in this class. While I might sometimes call on you during the class discussion, it is your responsibility to be an active participant without being called on.

Takehome Data Analysis Problem

The objective of the TakeHome is to give you independent applied research experience using real data and statistical methods. You will use all (relevant) techniques learned in this class to analyze a dataset provided by me.

Further details on the TakeHome will be provided as due dates approach.

Note that you **must score at least 30% of the points on the TakeHome Exam in order to pass this class**.

Exams:

There will be one midterm and one final in this class. See course info (/courses/Fall15/sta721/#exams) for dates and times of the exams. You are allowed to use one sheet of notes (`cheat sheet") on the midterm and the final. This sheet must be no larger than 8 1/2 x 11, and **must be prepared by you**. You may use both sides of the sheet and can write as small as you wish.

Email & Forum (Piazza):

I will regularly send announcements by email, please make sure to check your email daily.

Any non-personal questions related to the material covered in class, problem sets, labs, projects, etc. should be posted on Piazza forum (https://sakai.duke.edu/portal/site/ba0d1c18-ba55-473f-9d70-b6a1f9559bbe/page/08b2daf5-e24e-4401-97a1-ba4b8839895a#). Before posting a new question please make sure to check if your question has already been answered. The TAs and myself will be answering questions on the forum daily and all students are expected to answer questions as well. Please use informative titles for your posts.

Note that it is more efficient to answer most statistical questions ``in person" so make use of OH.

Students with disabilities:

Students with disabilities who believe they may need accommodations in this class are encouraged to contact the Student Disability Access Office (http://www.access.duke.edu/students/requesting/index.php) at (919) 668-1267 as soon as possible to better ensure that such accommodations can be made.

Academic integrity:

Duke University is a community dedicated to scholarship, leadership, and service and to the principles of honesty, fairness, respect, and accountability. Citizens of this community commit to reflect upon and uphold these principles in all academic and non-academic endeavors, and to protect and promote a culture of integrity. Cheating on exams and quizzes, plagiarism on homework assignments and projects, lying about an illness or absence and other forms of academic dishonesty are a breach of trust with classmates and faculty, violate the Duke Community Standard (http://www.studentaffairs.duke.edu/conduct/resources/dcs), and will not be tolerated. Such incidences will result in a 0 grade for all parties involved as well as being reported to the Office of Student Conduct (http://www.studentaffairs.duke.edu/conduct). Additionally, there may be penalties to your final class grade. Please review the Duke's Academic Dishonesty policies (http://www.studentaffairs.duke.edu/conduct/resources/academicdishonesty).

Policies:

- Late work policy for Homework:
  - next day: lose 30% of points
  - later than next day: lose all points
- Late work policy for TakeHome Data Analysis: 10% off for each day late.

- The final exam must be taken at the stated time.

- Regrade requests must be made **within 3 days** of when the assignment is returned, and must be submitted in writing. These will be honored if points were tallied incorrectly, or if you feel your answer is correct but it was marked wrong. No regrade will be made to alter the number of points deducted for a mistake. There will be no grade changes after the final exam.

- Use of disallowed materials (textbook, class notes, web references, any form of communication with classmates or other persons, etc.) during exams will not be tolerated. This will result in a 0 on the exam for all students involved, possible failure of the course, and will be reported to the Office of Student Conduct (http://www.studentaffairs.duke.edu/conduct). If you have any questions about whether something is or is not allowed, ask me beforehand.

▲