Course Information

Course Description

This is an introductory course in probability and statistics designed to develop a good understanding of uncertain phenomena and the mathematical tools used to model, analyze, and validate hypothesis.

Office Hours

Instructor:
Yunan Liu
Mudd 308a, Wednesday 2:00 - 4:00 pm
yl2342@columbia.edu

Teaching Assistants:
Yupeng Chen
Mudd 323, TBD
yc2561@columbia.edu
Xinyun Chen
Mudd 323, TBD
xc2177@columbia.edu

Recitation

Mudd 303, Friday 2:00 - 4:00pm

Prerequisites

A solid knowledge of calculus, including multiple variable integration.

Required Text

Ross, S. M. *Probability and Statistics for Engineers and Scientists.* Third Edition, Elsevier, 2004.

Homework

There will be weekly assignments due every Friday by 5:00 pm. There will be a box outside Mudd 306 to collect the homework, beware that the doors of the IEOR Department close at 5:00 pm. Students are encouraged to collaborate with other students in the class, as long as each person writes his/her own solutions. Copying homework from another student (past or present) is forbidden.

Graded assignments will be returned in class. Assignments not picked up in class will be placed in the IEOR 3600 homework box located outside of Mudd 306, where they will be kept only for two weeks. After two weeks any abandoned homework will be recycled.

Exams

1st midterm: Wednesday, February 25th, in class.
2nd midterm: Wednesday, April 1st, in class.
Final: TBD

Grading

Homework: 20%
1st midterm: 25%
2nd midterm: 25%
Final: 30%

Course Topics

1. Descriptive statistics
    - Describing data sets
    - Summarizing data sets
    - Chebyshevs inequality
    - Normal data sets
    - Paired data sets and the sample correlation coefficient
2. Elements of probability
    - Sample space and events
    - Venn diagrams and events
    - Axioms of probability
    - Conditional probability
    - Bayes Formula
    - Independent events
3. Random variables and expectations
    - Random variables
    - Types of random variables
    - Jointly distributed random variables
    - Expectation
    - Properties of the expected value
    - Variance
    - Covariance and variance of sums of random variables
    - Moment generating functions
    - Chebyshevs inequality and the Weak Law of Large Numbers
4. Special random variables

- Bernoulli and Binomial
- Poisson
- Hypergeometric
- Uniform
- Normal
- Exponential
- Distributions arising from the normal

5. Distributions of sampling statistics
   - The sample mean
   - The Central Limit Theorem
   - The sample variance
6. Parameter estimation
   - Maximum Likelihood Estimators
   - Interval estimates
   - Estimating the difference in means of two normal populations
   - Approximate confidence interval for the mean of a Bernoulli random variable
7. Hypothesis testing
   - Significance levels
   - Tests concerning the mean of a normal population
   - Testing the equality of means of two normal populations
   - Hypothesis tests concerning the variance of a normal population
   - Hypothesis tests in Bernoulli populations
8. Regression
   - Least squares estimators of the regression parameters
   - Distribution of estimators
   - Statistical inferences about the regression parameters
   - The coefficient of determination and the sample correlation coefficient
   - Analysis of residuals: assessing the model
   - Transforming to linearity
   - Weighted least squares
   - Polynomial regression
   - Logistic regression models for binary output data
9. Goodness of fit tests and categorical data analysis
   - Goodness of fit tests and categorical data analysis
   - Goodness of fit tests when some parameters are unspecified
   - Tests of independence in contingency tables