Department of Statistics

Syllabus for the First Year Exam (2020)

203: Introduction to Probability Theory

Textbook:

• DeGroot, M.H. and Schervish M.J. Probability and Statistics. Fourth Edition, Addison Wesley.

Course Topics:

- Sample space, sigma algebra, and the definition of probability.
- Counting methods. Combinatorial methods. Multinomial coefficients.
- Probability of a union of events. Conditional probability and independent events. Bayes' theorem.
- Probability density function, probability mass function, cumulative distribution function, and quantile function.
- Bivariate distributions and multivariate distributions.
- Conditional distributions and marginal distributions.
- Functions of a single random variable. Functions of two or more random variables. Probability integral transformation.
- Expectation and variance. Moments and moment generating functions. Covariance and correlation.
- Conditional arguments: conditional expectation, conditional variance, and condition distributions.
- Discrete random variables and their distribution properties. Examples: uniform, Bernoulli, binomial, multinomial, Poisson, hypergeometric, and negative binomial distributions.
- Continuous random variables and their distribution properties. Examples: uniform, normal, beta, gamma, and exponential distributions.

204: Introduction to Statistical Data Analysis

Textbook:

- R by Example (2012) by Jim Albert and Maria Rizzo, Springer Use R! Series

- Course slides available online.
- Other recommended books:
  - Modern Applied Statistics with S (2002, Fourth Edition) by W.N. Venables and B.D. Ripley, Springer.
  - R for Data Science (2017) by Garrett Grolemund and Hadley Wickham
  - ggplot2 (2016) by Hadley Wickham, Springer Use R! Series

Course Topics:

- Fundamentals of `R` programming.
- Numerical and exploratory data analysis in `R`.
- Data visualization and advanced graphs. Intro to `ggplot`.
- Hypothesis testing and linear regression models.
- ANOVA and randomized block models.
- Logistic regression.
- PCA, factor analysis and clustering.

205B: Intermediate Classical Inference

Textbook:

- Casella, G. and Berger, R. (2002). *Statistical Inference*, Second edition. Pacific Grove CA: Duxbury.

Course Topics:

- Sufficient and minimal sufficient statistics. Finding a sufficient statistic and showing that it is minimal sufficient. The factorization theorem. Examples from common distributions. Sufficiency and the exponential family. Ancillary statistics. Complete statistics. Techniques to find complete sufficient statistics. Examples mainly from exponential family. Basu's theorem connecting sufficiency, completeness and ancillary properties.
- Principles of statistical inference: Sufficiency, conditionality and likelihood.
- Methods of finding estimators: method of moments; maximum-likelihood estimation; unbiased estimators. Definitions and properties; examples with common distributions. Possible logical inconsistency of frequentist estimators (e.g., a negative estimate of a positive parameter).
- Methods for evaluating estimators: mean squared error, minimum variance, the Rao-Blackwell theorem, the Lehman-Scheffe theorem, the Cramer-Rao inequality.

- Methods for finding tests: likelihood-ratio tests.
- Methods for evaluating tests: error probabilities. Neyman-Pearson lemma for point null vs. point alternative. Extend it to composite null vs. composite alternative by using monotone likelihood ratio (MLR) property. UMP tests, p-values.
- Methods for finding interval estimators: Pivotal quantities, inverting a test statistic.
- Methods for evaluating interval estimators: coverage probability and size, illustrate in location and scale families how to find intervals with shortest length.
- Large sample theory for estimation, testing and confidence intervals.

206B: Intermediate Bayesian Inference

Textbook:

- Christian, R. (2001) *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation.* Springer, 2nd edition.
- Berger, J.O. (1984) *Statistical Decision Theory and Bayesian Analysis.* Springer.

Course Topics:

- Interpretations of probability: Subjective, Classical, Frequentist; Elicitation and representation of beliefs and uncertainty through probability distributions; Likelihoods.
- Priors and Posteriors: Bayes Theorem; Choice of priors; Calculation of posteriors; Conjugate analysis; Predictive distributions; Jeffreys priors and improper priors.
- Decision theoretic approaches to statistical inference; Expected losses; Frequentist and Bayesian risk; Optimality of Bayesian procedures.
- Exchangeability; Exponential families.
- Monte Carlo integration; Importance sampling; Laplace approximations.
- Markov chain Monte Carlo: Gibbs and Metropolis-Hasting sampling; Estimation of posterior distributions via MCMC; Drawing inferences, making predictions; Basic MCMC diagnostics; Simple hierarchical modeling.
- Bayesian Inference: Point estimation; Estimation Error; Interval estimation; Hypothesis testing; Bayes factors.
- Interpretation of results: Understanding posterior uncertainty.

207: Intermediate Bayesian Statistical Modeling

Textbook:

- Bayesian Data Analysis, Third Edition. A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari and D.B. Rubin. Chapman and Hall/CRC.

Course Topics:

- Single and multi-parameter models (Chapters 2 and 3)
- Normal approximations to the posterior distribution (Chapter 4)
- Bayesian inference for hierarchical models (Chapters 5 and 6)
- Model comparison and model assessment (Chapter 6)
- Modeling accounting for data collection (Chapter 7)
- Posterior simulation (Chapter 11-12)
- Approximations to the posterior distribution based on posterior modes (Chapter 13)
- Bayesian inference for regression models and hierarchical linear models (Chapters 14 and 15)
- Models for robust inference (Chapter 17)
- Mixture models (Chapter 22)

208: Linear Statistical Models

Textbooks:

- Christensen, R. Plane Answers to Complex Questions: Theory of Linear Models, Springer.
- Graybill, F. Theory and Application of the Linear Model, Duxbury.
- Searle, S. Linear Models, Wiley.
- Kutner, M., Nachtsheim, C., and Neter, J. Applied Linear Regression Models, McGraw-Hill.

Course Topics:

- Basic notions of linear algebra, e.g., vector spaces, column and null spaces of a matrix, inverse and generalized inverse, solutions to systems of linear equations, bases, orthogonal matrices, idempotent matrices, eigenvalues and eigenvectors.
- Definition and examples of the general linear model, including simple and multiple linear regression, analysis of variance, and analysis of covariance models.
- Ordinary and generalized Least Squares Estimation. Estimable functions. Best linear unbiased estimators and the Gauss-Markov Theorem.

- Distribution Theory. Class notes but available in many books. Covariances. Properties of covariances. Quadratic forms. Expectations of quadratic forms. Multivariate Normal distribution and its properties. Orthogonal transformations of MVN vectors. Partitions and conditional distributions. Quadratic forms in Multivariate Normal Variables and its distributions. Cochran's theorem. Non-central F distribution.

- Maximum likelihood estimation, interval estimation and hypothesis testing under the Gaussian Gauss-Markov model.

- You should also be familiar with fitting linear models using `R`. Class notes for examples.