List of Graduate Courses or Upper Division Mathematics/Statistics Classes

The following list will elaborate on textbook used and topics covered for all upper division and graduate statistics/mathematics course I have taken in my Undergraduate (coded as three digit course numbers). All my undergraduate courses are taken in UC Berkeley's Statistics or Mathematics Department.

Stats 210A: Theoretical Statistics (First Year Ph.D. Class, In Progress) **Professor: Will Fithian**
Text: Keener, Theoretical Statistics: Topics for a Core Course, Springer 2010.

We start this class off with a brief mention of measure theoretic concepts, although this will not be the emphasis of this course. We then talk about estimation, risk functions, sufficiency principle, Rao-Blackwell Theorem, and Cramer-Rao lower bound. We then move onto a more Bayesian framework but still with a Frequentist motivation behind why such a framework is attractive (i.e. minimizing average risk). We talk about Bayesian estimators, Empirical Bayes, hierarchical model, James-Stein paradox, and minimax estimators. We move back to the Frequentist setting to talk about hypothesis testing and Neyman-Pearson Lemma, UMP, UMPU, linear models, probability convergence, asymptotic theory, and MLE. We are currently talking bootstrap and the theory behind it, and if we have time we will also talk about multiple testing.

Because this is not only my first Ph.D. course but also an extremely difficult one, I have spent a tremendous amount of time doing the weekly problem sets. For those interested in seeing the most recent statistics course I have been doing please go to the following link under "Assignments" to see the weekly problem sets.
https://www.stat.berkeley.edu/~wfithian/courses/stat210a/

Stats 154 - Machine Learning (In Progress) Professor: Tuhao Chen
Text: An introduction to Statistical Learning (ISL) by James, Witten, Hastie, and Tibshirani Springer, 2013,
The Elements of Statistical Learning (ESL) by Hastie, Tibshirani and Friedman. Springer, 2009 (2nd Ed).

We begin by talking about the framework of machine learning then begin talking about supervised machine learning with linear regression. We then talk about classification methods that move beyond linear problems and in the framework of Bayesian analysis (discriminant analysis), KNN, regression trees, ridge and LASSO regression, and different penalty criterias. We also then talk about boosting and additive models, resampling methods, random forests and Kernel methods support vector machines. Currently we will now talk about unsupervised learning dealing with the "curse of dimensionality" problem with principle component analysis and cluster analysis.

Stats 153 - Time Series (A) Professor: Alexander D'Amour
Text: Time Series Analysis and Its Applications by R.H. Shumway and D.S. Stoffer

We start off with learning different transformation, smoothing, and identifying autocorrelations via visual plots. Then we talk about stationary and linear process, while building a theoretical framework using backshift polynomials and k^{th} order differencing to ultimately build ARMA models. Eventually we talk about the theory behind ARMA, which leads to SARIMA models, including discussion about Yule-Walker equations, Barlett's formula, and some asymptotic distributional remarks. We then, for the sake of the project also, talk about model selection theory using machine learning techniques such as cross validation, AIC/BIC, Mallow's Cp. The last part of this class included talking about the frequency domain, and breaking up the model into its spectral density while looking it through the power transform function.

The project required us to make predictions about Google Query time series trends for certain keyword searches. Using mainly SARIMA models, we used model selection theory about which parameters of the SARIMA model gives the best score function using cross validation or BIC. The candidate models were chosen from ACF and PACF plots, which gives an idea of what kind of SARIMA parameters we should consider. Seasonal differencing was used to also transform the data into a stationary model to justify using a SARIMA model.
Note: All projects are available on my github link: https://github.com/DavidHam97/Undergraduate

Stats 151A - Linear Modeling (A+) Professor: Shobhana Stoyanov
Text: Applied Regression Analysis and Generalized Linear Models, J. Fox

This class starts off with talking about transformations of data to make a non-linear relationship to a linear relationship along with a discussion of the loss of interpretation through such a method. We then talk about Linear Least Square Regression through the approach of calculus optimization then a different approach through linear algebra vector geometry with orthogonal projections and matrix manipulation. We then derive the distributional assumptions used for normal linear least square regression such as explaining how the normality of errors give us the required distributions of the coefficients being t-distributed and ultimately talk about the Gauss-Markov Theorem and the detailed proof of it. We then talk about regression diagnostics to check whether crucial assumptions for the linear model holds. Through this discussion we talked about the robustness of the models such as multicollinearity problem and how principle component analysis, a dimension reduction technique, is used to deal with such problems. At the latter part of the course. We focused on logistic regression, dummy variables, interaction effects, and bootstrapping regression models.

There was also an individual project in the latter part of this class where I analyzed forest fire data, in which I had to predict whether or not given certain covariates (around 10 covariates were available) there would be a fire in a certain grid location in the *xy* plane. This involved choosing the right covariates, i.e. model selection, explaining whether or not such a linear relationship was even appropriate after certain transformations, i.e. regression diagnosis, and finally to interpret the usefulness of my results in the context of the problem.

Stats 152 - Sampling Surveys (A+) Professor: Shobhana Stoyanov
Text: Sampling, S. Lohr (2nd Edition)

We first talk about what makes a good sample, what are some possible bias, and how to identify and fix those. We then talk about the theory of sample inference. We begin by talking about simple random sample and then talk about stratified sampling, cluster sampling, one-stage and two-stage stratified cluster sampling, then complex surveys which uses a combination of all of these. We then talk about sampling with unequal probabilities such as proportional sampling which can often give smaller variances according to Horvitz-Thompson estimator. We also discuss how to estimate variances for much more complex surveys such as bootstrapping, replicated repeating sampling, jackknifing, imputation, and linearization methods using Taylor's theorem. Lastly we talk about categorical data analysis in the context of a Chi-square test. This class was extremely useful to those like me interested in social science research where mainly the sample of study are people.

A big portion of this class was project based. My group decided to analysis the 2016 National Crime Victimization Survey (NCVS) that is publicly available on the government website to make inference and possible associations of crime statistics with certain categories such as race, gender, education level, etc. If there were strong correlations, we study possible socio-economic reasons for such a circumstance. The theory that was involved in this project was dealing with nonresponse, making careful variance calculations, doing Chi-square tests of associations, and finally breaking the "complex survey" into a manageable dataset.

Stats 155 - Game Theory (A) Professor: Shobhana Stoyanov
Text: Game Theory, Alive by Anna R. Karlin and Yuval Peres ,Game Theory by Thomas S. Ferguson

This was a quite a heavy theoretical applied math class that started off with learning combinatorial games such as Nim, Hex, etc. We move on to talk about zero-sum and general-sum games where we explored payoff matrices, Nash Equilibria, and the proof of Neumann's Minimax Theorem. The next unit was about evolutionary and correlated equilibria especially in applications of evolutionary biology and touched briefly on stable matching and allocations, covering the Gale-Shapley algorithm. We lastly talk about auctions where probability densities were introduced to figure out the optimal payoff for assuming certain distributions on the bidder.

Stats 135 - Concepts of Mathematical Statistics (A+) Professor: Herman Pitters
Text: J. A. Rice. Mathematical Statistics and Data Analysis. 3rd Edition.

We begin this mathematical statistics course talking about survey estimation theorem and eventually talking about estimation of parameters in distributions. The methods we explore are method of moments, maximum likelihood estimators, and the Bayesian approach to estimation. We do talk a bit about sufficiency and the Rao-Blackwell theorem, but without proof. We move onto hypothesis testing and look at the Neyman-Pearson Paradigm of a frequentist approach to understanding the concept of a p-value. We talk about likelihood ratio tests, uniformly most powerful tests, and the duality between confidence intervals and hypothesis testing. We then move onto talking about bootstrapping and building empirical CDFs for non-parametric estimation. We also talk about comparing two samples using hypothesis testing such as the signed rank test and the Mann-Whitney test.

Stats 134 - Concepts of Probability (A) Professor: Ani Adhikari
Text: Probability, by Jim Pitman.

We cover almost the whole book where we first talk about basic probability rules, distributions, and definition and properties of expectation. We then move on to define and explore the properties of certain discrete distributions such as binomial and Poisson distributions, studying their moments and variances. We move onto the continuous distribution case where the measure is different, dealing with integrals now to find probabilities. We talk about the exponential and gamma distribution, the memoryless property and the change of variable formula for specific transformations of the distribution. We also talk about order statistics, finding their distributions, Chebyshev's inequality, and the convolution method.

Stats 133 – Concepts in Computing with Data (A-) Professor: Deborah Nolan
Text: Automated Data Collection with R: A Practical Guide to Web Scraping and Text Mining by Christian Rubba, Dominic Nyhuis, and Simon Munzert

This class begins by studying the basics of R including sub setting and data frame manipulation. We then learn data visualization (using the package ggplot) and data scraping. We then learn how to build functions and other basic built in functions for easier looping such as sapply(). We also learn regular expressions, basic SQL statements in R, and HTML.

This class was heavily project based. In the first project I analyzed node configurations. Some of the skillsets used were linear algebra eigenvalue decomposition, numerical approximations such as bisection method, matrix computation in R, and random number generation. In my second project I analyzed the election between Trump and Clinton from a county level data. We tried classifying patterns within race, education, gender in voting behavior and created visual map plots to summarize the information. We then tried to predict the 2016 election results from previous election data using cross validation and classification methods.

Math 113 - Abstract Algebra (In Progress) Prof: Vivek Shende
Text: Algebra by Artin (2^{nd} Edition)

We begin by talking about some basic number theory claims from discrete mathematics. We then talk about equivalence relations and how the equivalence classes partition the set given an equivalence relation. After this we dive into group theory, which includes subgroup, homomorphism, cosets, Lagrange Theorem, Correspondence Theorem, First Isomorphism Theorem, and Isometries. We then talk about dihedral groups, orbit stabilizer formula, permutations, plane crystallographic groups, conjugation, and Sylow theorems. In the second part of this class we move onto ring and field theory. We redefine previous definitions such as homomorphism in this setting and talk about imaginary number field, factoring polynomials, Gauss' Lemma, and field extension.

Math 104 - Real Analysis (A) Prof: David Dynerman
Text: Principles of Mathematical Analysis by Walter Rudin

We went sequentially in order from chapter 1 to 7 of Rudin. These topics include sets and real number system, basic topology, including ideas of compactness, openness, and connectedness, and metric spaces. We then talked about sequences and series, which include various topics such as Cauchy sequences, with examples to illustrate convergence and absolute convergence. Chapter 4 was about continuity, where limits were introduced, compactness and continuity properties of functions. We then talked about differentiation in Chapter 5, and we also proved many useful applications such as Mean Value Theorem, L'Hospital's rule and Taylor's Theorem. Chapter 6 was about the Riemann-Stieltjes integral and the properties and existence of this integral, covering the rectifiable curve unit also. Finally, we ended with Chapter 7 on sequences of functions talking about uniform convergence in the perspective of almost everything we learned for example in continuity, integration, and differentiation, covering also lastly the Stone-Weierstress Theorem.

Math 185 - Complex Analysis (A) Prof: Tim Laux
Text: Introduction to complex variables and applications by Ruel Vance Churchill

We covered sequentially from chapter 1 to 8 of this textbook. We began with review of basic analysis tools such as definition of limit, epsilon delta definitions, then talked about exponential and polar form of complex numbers, analytic functions and Cauchy-Riemann equations, Harmonic functions, and sufficient criteria's for differentiability. We then talk about elementary complex functions such as trigonometric and logarithmic functions defined on different branches and finally move on to integrals of complex functions. These involved the Cauchy-Goursat Theorem, Cauchy Integral Formula, Liouville's theorem and fundamental theorem of Algebra while exploring also the Maximum Modulus Principle. We then learn the Laurent's and Taylor series for the complex plane while also learning Residues and applications of Residues in finding integrals using these Laurent series expansions.

Math 110 - Linear Algebra (A) Prof: Zvezdelina Stankova
Text: Linear Algebra by Arnold J. Insel, Lawrence E. Spence, and Stephen H. Friedberg

We managed to cover the whole book. Chapter 1 briefly covered basic definitions of vector spaces, since it was covered extensively on previous elementary linear algebra class. We then talked about linear transformations, invertibility, isomorphism, equivalence relations, change in coordinate basis, null spaces and range, and very briefly talked about determinants and solving for systems of linear equations. We then spent in great detail the sufficient and necessary conditions for diagonalization, leading to studying also the Jordan canonical form. The last topic was on inner product, touching on applications with Gram-Schmidt Process, Adjoints, Normal operators, Unitary and orthogonal projections, with application to linear regression

Math 128A - Numerical Analysis (A) Prof: Ming Gu
Text: Numerical Analysis by Richard L. Burden and J. Douglas Faires

This was both a theory and programming class with Matlab. We covered chapters 1 to 6. This includes numerical approximation methods for finding zeros such as bisection method, Newton's method, error analysis for order of convergence. We then covered interpolation and polynomial approximations using Lagrange polynomials, cubic spline interpolation and Hermit interpolation in Chapter 3. Chapter 4 was about numerical differentiation and integration using Richardon's extrapolation and Quadrature methods. We then talked about approximating ordinary differential equations using Runge-Kutta methods, multistep-methods and variable step-size methods in chapter 5. Lastly, in Chapter 6, we focused on theory in matrix algebra and matrix factorization such as the LU factorization and Cholseky factorization.

The project focused on computer programming which was to numerically approximate the ordinary differential equations for gravity and our solar system's orbits for specifically Earth, Moon and the Sun. Our final task was to create a plot that accurately showed the rotation and derive the numerical approximations for those ordinary differential equation solutions.