Syllabus for Probability & Statistics Review

Course Section: Probability & Statistics, ECON 508B, Summer 2020
Time: 10:00 AM-12:00 PM (Mon.-Fri.) Aug. 24^{th}-Sep. 11^{st}, 2020
Instructor: Hongyi Liu
Email: hongyi.liu@wustl.edu
Office: 354 Seigle Hall
Office Hours: 10:00 AM-12:00 PM, Saturday

Course Description

This course is a review of the fundamental concepts of econometrics, the basic mathematical tools of probability theory, measure theory and statistical theory used in the core courses of the 1^{st} year doctoral graduate program in Economics (& Business school) at Washington University in St. Louis. It aims to cover some of prerequisites for *Introduction to Econometrics, Applied Econometrics, & Quantitative Methods in Economics*, of which also further stimulate the deep understanding of some of Micro and Macro models. **In addition, this year I intend to incorporate a week-series review of the state-of-the-art methodologies in *Machine Learning, Deep Learning, and Natural Language Processing (NLP).*** It time permits, we can also go through some classical algorithms and data structure frequently used in Machine Learning via *Python*, such as BFS, DFS, Binary search and binary tree, Two pointers, Hash & Heap, dynamic programming, etc. Basic knowledge about string, array list, queue, and stack will be helpful to this section. (C++ and Java are also fine in this part).

Materials & References

Lecture Notes:

Lecture notes and slides for this math camp can be found on the Ph.D. (ECON) program description page. Note that some of slides could be updated as we go.

References:

This course is intended to be self-contained, but you may find the following recommended textbooks useful and worthwhile, especially if you have limited probability and statistics experience.

Introductory Econometrics:

1. Hansen, Bruce E., *ECONOMETRICS*. University of Wisconsin, *manuscript*.
2. Wooldridge, Jeffrey M., *Introductory econometrics: A modern approach*. Nelson Education, 2015.
3. Hayashi, Fumio, *Econometrics*. New Jersey, USA: Princeton University, 2000.

4. Stock, James H and Watson, Mark W, *Introduction to Econometrics*. Pearson Education Boston, MA, 2012.

5. Hsiao, Cheng, *Analysis of panel data*. Cambridge university press, 2014.

Probability Theory & Measure Theory:

1. Ash, Robert B., and Catherine Doleans-Dade. *Probability and measure theory*. Academic Press, 2000.

2. Pollard, David. *A user's guide to measure theoretic probability*. Vol. 8. Cambridge University Press, 2002.

3. Athreya, Krishna B., and Soumendra N. Lahiri. *Measure theory and probability theory*. Springer Science & Business Media, 2006.

4. Tao, Terence, ed. *An introduction to measure theory*. Vol. 126. American Mathematical Soc., 2011.

5. Stein, Elias M., and Rami Shakarchi. *Real analysis: measure theory, integration, and Hilbert spaces*. Princeton University Press, 2009.

6. Wheeden, Richard L. *Measure and integral: an introduction to real analysis*. Vol. 308. CRC press, 2015.

Statistical Theory:

1. Casella, George, and Roger L. Berger. *Statistical inference*. Vol. 2. Pacific Grove, CA: Duxbury, 2002.

2. Hogg, Robert V., and J. W. McKean. Allen T. Craig. *Introduction to mathematical Statistics*. 2005

3. Lehmann, Erich Leo, and George Casella. *Theory of point estimation*. Springer Science & Business Media, 2006.

4. Van der Vaart, Aad W. *Asymptotic statistics*. Vol. 3. Cambridge university press, 2000.

Grading Allotment

- In-Class Attendance...........................50%.
- Homework...30%.
- In-Class Exam (Or Take-Home).........20%.

Tentative Course Calendar

- Lecture 0: Introductory Econometrics
  - Conditional Expectation and Projection, The Algebra of Least Squares, Least Squares Regression, Time Series Regression, Endogeneity (IV), Correlation v.s. Causal effect, Large Sample Asymptotics, Panel Data Model, Machine Learning.
- **Lecture X: Machine Learning, Deep Learning, NLP**
  - Supervised Learning, Basis Expansions and Regularization, Kernel Smoothing and Model Assessment, Model average, Trees, Boosting, Neural Networks, SVM, Random Forests, CNN, RNN, LSTM, Autoencoders, Unsupervised Learning, Ensemble Learning, Bootstrap, Monte Carlo, Word2Vec, fastText, GloVe, and Numerical Computation.
- Lecture 1: Probability and Distribution
  - Set theory, Fundamentals of Measure and Probability Set Function, Conditional Probability and Independence, Induced Measure, Random Variables and Distribution Functions.
- Lecture 2: Lebesgue Measure and Lebesgue Measurable Functions
  - Review of Real Analysis, Lebesgue Outer Measure, Lebesgue Measure, Lebesgue Measurable Functions.
- Lecture 3: Riemann, Lebesgue Integral and Repeated Integration
  - Review of Riemann Integral, Lebesgue Integral, Comparison between Lebesgue and Riemann Integral, Repeated integration
- Lecture 4: Modes of Convergence and Important Inequalities
  - Convergence for Measurable Functions, Important Inequalities
- Lecture 5: Expectation, MGF and CGF
  - Expected Values, Moment Generating Functions, Cumulative Generating Functions
- Lecture 6: Statistical Theory
  - Parametric Families of Distributions, Decision Theory, Data Reduction
- Lecture 7: Optimal Data Reduction
  - Exponential Families, Minimal Sufficiency, Completeness, Risk Reduction, Optimal Unbiased Estimation
- Lecture 8: Model Estimation

  - Bayesian Estimation, MLE, GMM

- Lecture 9: Stochastic Convergence and Orders
  - Basics of Stochastic Convergence, Convergence in Distribution, Orders of Magnitude, Stochastic Orders of Magnitude

- Lecture 10: LLN, CLT and Local Linear Approximation
  - Laws of Large Numbers, Continuous Mapping Theorem, Linderberg-Feller Central Limit Theorem, Delta Method