MECO 7312: Advanced Statistics and Probability

Dr. Khai Chiong

Fall, 2020

E-mail: khai.chiong@utdallas.edu
Office Hours: by appointment
Office: JSOM 13.324
Web: **www.khaichiong.com/teaching**
Class Hours: Wednesday 1-3:45pm
Class Room: JSOM 2.722

Course Modality

The modality for this course will be an in-person, traditional classroom format. However, students can choose to participate in this course remotely, either synchronously or asynchronously. Lectures will be streamed live via Microsoft Teams, where you can participate and ask questions synchronously. In addition, lectures will be recorded and will be available asynchronously.

Course Description

The goal of this PhD-level course is to develop a rigorous theoretical foundation necessary for research in applied econometrics and statistics. It covers the core topics of probability theory and statistical inference, including properties of random variables and probability distributions, frequentist and bayesian estimation, asymptotic theory, and various methods of hypothesis testing.

Texbooks

The main required textbook is *Statistical Inference* by Casella and Berger.

- Casella, George, and Roger L. Berger. *Statistical inference*. 2nd Edition
- Davidson, Russell, and James G. MacKinnon. *Econometric theory and methods*.
- Cameron, A. Colin, and Pravin K. Trivedi. *Microeconometrics: methods and applications.*

Grades

Course grade is based on the weighted average of four problem sets and one final exam. Please work together in a team of 4 to 5 for the problem sets – we often learn a great deal from our peers.

Computing

You should be familiar with basic statistical programming in at least one programming language. I use the following languages: R, Python, MATLAB, STATA, Mathematica. With the exception of Stata, all of the above are free for UTD graduate students. In addition, you are expected to learn and use LaTeX for typesetting your assignment.

Schedule

The schedule is tentative and subject to change.

Week 1: Basic probability theory (Casella-Berger, chapter 1): sample spaces, event spaces, probability spaces, random variables, probability density functions, cumulative density functions.

Week 2: Transformations and expectations of random variables (Casella-Berger, chapters 2): change of variables, probability integral transformation moments and expectations.

Week 3: Multivariate random variables (random vectors) (Casella-Berger, chapter 4): joint and marginal distributions, conditional distributions, independence of random variables; covariance and correlation, bivariate transformations

Week 4: Common families of distributions (Casella-Berger, chapters 3): Multivariate Normal distribution, Gamma distribution, truncated random variables, probability inequalities.

Week 5: Properties of a random sample (Casella-Berger, chapter 5): sampling distributions, order statistics, unbiasedness and consistency, convergence concepts (convergence in probability, convergence almost surely, convergence in distribution).

Week 6: Point estimation (Casella-Berger, chapter 5): large sample theory, central limit theorem, delta method, big o notation, asymptotic variance

Week 7: Point estimation (Casella-Berger, chapter 7): Method of Moments estimator, Generalized Method of Moments, Maximum Likelihood Estimation

Week 8: (Point estimation (Casella-Berger, chapter 7): properties of the Maximum Likelihood estimator, loss functions, mean square error, Fisher's information, Probit models.

Week 9: (Casella-Berger, chapter 7. Cameron-Trivedi, Chapter 13): Bayesian methods versus Frequentist. Conjugate prior. Bayes Theorem. Bayesian estimation.

Week 10: (Casella-Berger, chapter 8): Hypothesis testing, Likelihood Ratio test, Wald's t-test, Type-1 and Type-2 errors.

Week 11: (Casella-Berger, chapter 8): Lagrange-multiplier test. Size and power of a test. p-values. Neyman-Pearson lemma. Asymptotic distribution of test statistics.

Week 12: (Casella-Berger, chapter 9): Confidence interval as inversion of test statistics. Coverage probabilities. Bayesian intervals.

Week 13: (Casella-Berger, appendix): Data-resampling and simulation techniques. Bootstrapping. Importance sampling. Monte Carlo sampling and integration. Non-parametric methods. Kernel estimators. k-nearest neighbors.

Week 14: (Davidson-Mackinnon, chapters 2-3): Statistical properties of linear regressions. Matrix notation. Bias and consistency. Frisch-Waugh-Lovell Theorem. The geometry of OLS estimation. Multicollinearity.

Week 15: (Davidson-Mackinnon, chapters 2-5): OLS covariance matrix. Hypothesis testing and inference involving OLS estimators. Heteroskedasticity consistent covariance matrix estimator. Serial correlation.

Optional class 1: Causal inference methods. Average treatment effect. Propensity score matching. Difference-in-difference methods.

Optional class 2: Discrete-choice models. Probit and Logit models. Binary and multinomial models. Time-series and panel data. Testing for serial correlations. Random and fixed effects.

Overlap with other courses

The first part of this course (the *probability* part of this course) overlaps with OPRE 7310 Probability and Stochastic Processes. The treatment of probability is more rigorous and measure-theoretic than the treatment here.