BSTA 622 Statistical Inference II
Fall 2023

Content:
This course focuses on theoretical statistics. We will cover a series of classical statistical inferential methods, including the asymptotic theory for maximum likelihood estimation, the method of estimating equations, the generalized method of moment estimation, and inference by influence functions. This course will emphasize concepts, methods and theories, rather than applications. Successful completion of this course will provide you with a foundation in probability-based statistical inference.

Intended Audience:
The course is designed for Biostatistics Ph.D. students in their 2nd year or beyond. Students are required to complete Probability I (BSTA 620) and Inference I (BSTA 621) before taking this course. Exceptions may be made with permission of the instructor.

Instructors:
Jing Huang, PhD
Jing14@pennmedicine.upenn.edu
Jinbo Chen, PhD
jinboche@pennmedicine.upenn.edu

TA:
Dazheng Zhang
Dazheng.Zhang@pennmedicine.upenn.edu
Office Hours: TBD

Class Schedule:
Mon and Wed 10:15-11:45 am in Blockley Hall 1311.

Textbooks:
Recommended, not required, textbooks:
Theory of Point Estimation, by E.L. Lehmann and G. Casella, Springer
Elements of Large-Sample Theory, by E.L. Lehmann, Springer
Asymptotic Statistics, by A.W. van der Vaart, Cambridge
Theoretical Statistics, by D. Cox and D. Hinkley, Chapman and Hall

Grading:
The goal of this course is to introduce students to classic topics in statistical inference, familiarize them with seminal literature, inspire deep discussions to understand the theories, and equip them with foundational skills for theoretical research. Therefore, this course will emphasize understanding over examination. Grades will be determined by class interactions, discussions, and presentations.

Paper reading and presentation:
In addition to lectures, there will be four "paper reading and discussion" sections in this course. Each section will closely align with the topics covered in the class materials. Each section consists of two classes: one for paper reading and the other for paper presentation.

Paper Reading Class:
Only students and the TA will participate in the paper reading session. Students should arrive having already read the paper and formulated questions. They should also research related literature, either to aid their understanding of the paper or to broaden their grasp of the topic, prior to attending the paper reading class. This session will be an opportunity for them to discuss, pose questions, and engage in discussions with their peers.

Paper Presentation:

After the reading class, students will decide on the presentation format for the paper. During the presentation class, half the students will collaboratively present the paper. Every student is expected to participate actively in speaking. The presentation should be comprehensive, covering the paper's motivation, intuitive ideas, mathematical aspects, implications, and potential extensions.

Your final grade will be influenced by your performance in the paper presentation class and the TA's feedback on your participation during the paper reading class.

Tentative Schedule

Date		Topics
Aug	30	Mathematics Primer
Sep	4	Labor Day no class
	6	Mathematics Primer
	11	Large Sample Theory
	13	Asymptotic Properties of the MLE (normality)
	18	Asymptotic Properties of the MLE (Efficiency)
	20	Asymptotic Properties of the MLE (Consistency)
	25	Paper reading: Self, S.G. and Liang, K.Y., 1987. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, *82*(398), pp.605-610.
	27	Asymptotic Theory of Estimation
Oct	2	Influence Functions
	4	Paper presentation Q&D for all the materials covered so far
	9	Statistical Information
	11	Statistical Information
	16	Paper reading: Reid, N. and Fraser, D.A.S. (2003). Likelihood Inference in the Presence of Nuisance Parameters. *PHYSTAT2003, SLAC*, 265-271 (arXiv:physics/0312079).
	18	Generalized Linear models
	23	Paper presentation Q&D for all the materials covered so far
	25	Hypothesis testing: Likelihood ratio test, Score test, and Wald test
	30	Profile likelihood estimation
Nov	1	Instructor-led paper discussion: Chatterjee N, Carroll R, Semiparametric maximum likelihood estimation exploiting gene-environment independence in case-control studies, *Biometrika*, Volume 92, Issue 2, June 2005, Pages 399–418
	6	Unbiased estimating functions/Estimation under model mis-specification
	8	Unbiased estimating functions/Generalized Method of Moments
	13	Pseudo-likelihood method
	15	Instructor-led paper discussion: Che, M., Han, P. & Lawless, J. F. (2023). Improving estimation efficiency for two-phase, outcome-dependent sampling studies. Electronic Journal of Statistics 17, 1043–1073.
	20	Paper reading: Gronsbell J, Liu M, Tian L, Cai T. Efficient Evaluation of Prediction Rules in Semi-Supervised Settings under Stratified Sampling. J R Stat Soc Series B Stat Methodol. 2022; ;84(4):1353-1391.
	22	Thanksgiving No class
	27	Paper presentation Q&D for all the materials covered so far
	29	Empirical Likelihood Estimation (Reference papers: Owen 1988 Biometrika, Qin 1999 Biometrika, Qin & Zhang 2005 Biometrika)

Dec	4	Empirical likelihood estimation
	6	Paper reading: Prosenjit Kundu and others, Generalized meta-analysis for multiple regression models across studies with disparate covariate information, *Biometrika*, 2019, Pages 567–585
	11	Paper presentation Q&D for all the materials covered so far