Stats 201C Advanced Modeling and Inference

Course site on CCLE: https://ccle.ucla.edu/course/view/19S-STATS201C-1.
Instructor: Qing Zhou (zhou@stat.ucla.edu), OH: Thursday 2pm–3pm, MS 8979.
TA: Hao Wang (hwang16@ucla.edu), OH: Tuesday 3pm–5pm, MS 8145.
Prerequisite: Stats 200B and 201B (recommended).
Programming skills (R, C/C++, Matlab, etc.).

Grading
Final grades of this course consist of three parts:
1. Homework assignments (40%): five homework assignments.
2. Midterm exam (30%): Open-book, date TBD.
3. Final paper (30%): due final exam week.

Course description
Introduction to advanced topics in statistical modeling and inference. The course covers three
groups of topics:
• Incomplete data and hidden variable models;
• Network data and random graphs;
• Graphical models and causal networks.

The course also introduces computational methods developed for these models and problems.

List of topics
Below is a tentative structure of the course:
1. Incomplete data and the EM algorithm: assumptions of missing data, EM and its proper-
ties, incomplete multivariate normal data.
2. Hidden variable models: mixture modeling, EM clustering, hidden Markov models (HMMs),
prediction and estimation for HMMs.
3. Network data: stochastic block models, variational EM, exchangeable random graphs, 
graphon.
4. Graphical models: undirected graphs, Gaussian graphic models, directed acyclic graphs
(DAGs), Bayesian networks, conditional independence, graph separation.
5. Causal DAGs: experimental intervention, identification of causal effects, structural equa-
tion models, counterfactuals, structure learning of DAGs.
References

• Lecture notes: to be posted on the CCLE site weekly.
• Schafer, J.L., Analysis of Incomplete Multivariate Data (1997).