SYLLABUS

USEFUL BOOKS

- BDA3 2014
- Wooldridge 2010
- Goodfellow et al (Deep neural networks)2016

1. APPROACHES TO INFERENCE WITH LOTS OF DATA

Simple models, weak assumptions, reliance on asymptotic distribution theory: With big data sets, the assumption that the sample size is "large" is reliable, so we can use asymptotic distribution theory confidently — so long as we are not tempted into constructing models whose parameter space grows with sample size.

Explicit probability distribution for the data, likelihood-based inference: With fast computers and Markov Chain Monte Carlo and related methods, we can construct and reliably estimate more believable (and hence usually more complicated) models without relying on asymptotic approximations.

Expanding models: Sieve estimation. Bayesian infinite-dimensional parameter spaces. Combine explicit probability models with (asymptotically) weak assumptions.

Readings.

- Angrist and Pischke 2010
- Sims comment on them. `http://sims.princeton.edu/yftp/AngristPischkeJEP/AngristPischkeComment.pdf`
- Sims "Understanding Non-Bayesians" 2011
- Mueller-Norets bettable CI's.2016

2. APPLYING THE APPROACHES TO BIVARIATE REGRESSION

OLS, sandwich, GLS: Sequentially relaxing assumptions leads asymptotically from OLS to sandwich?

Modeling the distribution of residuals: Do semiparametric efficiency bounds imply this is pointless?

Nonlinear regression: Asymptotically, a conscientious econometrician probably never continues to maintain a linearity assumption

3. MCMC

4. MULTI-LEVEL MODELS

- Why Bayesian approaches are natural for these models.

- Hierarchical priors in general
- "Fixed effects", "random effects", "correlated random effects"
- Clustered standard errors vs. modeling heteroskedasticity.
- Weighted data.
- Dynamic panel models
- Panel VAR's

Readings.

- Wooldridge book.
- BDA3
- Laura Liu, "Density Forecasts in Panel Data Models: A Semiparametric Bayesian Perspective". `http://www.princeton.edu/~erp/erp%20seminar%20pdfs/panel_dfcst_paper_v10.1%20(Laura%20Liu).pdf`
- Skinner and Wakefield on survey data (with weights), 2017

5. Regression with many regressors; IV with many instruments.

- LASSO, double-exponential prior
- $N(0, \lambda I)$ prior, ridge
- $\theta^{length(\beta)}$ prior
- Principal components to shrink X space. Factor Analysis.

6. Many parameters plus nonlinearity

- Gaussian process priors
- Deep neural nets.

Readings.

- Goodfellow book
- Sims Gaussian process prior example. 2000

References

Angrist, J. D., and J.-S. Pischke (2010): "The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics," *Journal of Economic Perspectives*, 24(2), 3–30.

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin (2014): *Bayesian Data Analysis*. CRC Press.

Goodfellow, I., Y. Bengio, and A. Courville (2016): *Deep Learning*. MIT Press.

Müller, U. K., and A. Norets (2016): "Credibility of Confidence Sets in Nonstandard Econometric Problems," *Econometrica*, 84(6), 21832213.

Sims, C. A. (2000): "Using a Likelihood Perspective to Sharpen Econometric Discourse: Three Examples," *Journal of Econometrics*, 95(2), 443–462, `http://www.princeton.edu/~sims/`.

——— (2011): "Understanding non-Bayesians," Discussion paper, Princeton University, http://sims.princeton.edu/yftp/UndrstndgNnBsns/GewekeBookChpter.pdf.

SKINNER, C., AND J. WAKEFIELD (2017): "Introduction to the Design and Analysis of Complex Survey Data," *Statistical Science*, 32(2), 165–175.

WOOLDRIDGE, J. M. (2010): *Econometric Analysis of Cross-Section and Panel Data*. MIT, second edn.