

Reference Materials for STATS 551

Yang Chen, Department of Statistics, University of Michigan

1 Review of Probability and Statistical Inference

- Expectation & Variance.
- Basic Distributions (Gaussian, Binomial, Beta, Dirichlet, Gamma, Wishart).
- Likelihood. Confidence Intervals.

Major references:

- Appendix [15]. Chapter 2 of [9].
- Lecture notes ‘Introduction.pdf’.
- Exercise: Homework 0.

2 Introduction to Bayesian Inference

- Bayes Formula.
- Prior. Likelihood. Posterior.

Major references:

- Chapter 1 of [7]. Chapter 1 of [9].
- Lecture notes ‘BasicsofBayesInference.pdf’.
- Andrew Gelman’s blog post: <http://andrewgelman.com/2016/12/13/bayesian-statistics-whats/>

3 Single Parameter Models

- Estimating single mean.
- Estimating single variance.
- Estimating probability.
- Choices of priors.

Major references:

- Chapter 2 of [7]. Chapters 3, 4, 5 of [9].
- Lecture notes ‘SingleParameterModels.pdf’.
- Exercise: Homework 1.

4 Multi-parameter Models

- Multinomial models.
- Multivariate normal models.
- Nuisance parameters.

Major references:

- Chapter 3 of [7]. Chapter 7 of [9].
- Lecture notes ‘MultiparameterModels.pdf’.

5 Hierarchical models

- Conjugate models.
- Hierarchical models.
- Shrinkage estimators.

Major references:

- Chapter 5 of [7]. Chapter 8 of [9].
- Lecture notes ‘HierarchicalModels.pdf’.
- Exercise: Homework 2.

6 Bayesian Computation

- Importance sampling
- Gibbs sampling
- Metropolis-Hastings algorithm
- Hamiltonian Monte Carlo

Major references:

- Chapters 10, 11, 12 of [7]. Chapter 6, 10 of [9]. Chapters 5, 6,9 of [12].
- Lecture notes ‘BayesComputing.pdf’.
- Illustrations for Metropolis-Hastings and Hamiltonian Monte Carlo [14].
- Exercise: Homework 3.

7 STAN

R STAN getting started: <https://github.com/stan-dev/rstan/wiki/RStan-Getting-Started>

Prior choice: <https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations>

Example models: <https://github.com/stan-dev/example-models/wiki>

Major references: [6, 8, 7]

8 Regression Models

- Linear regression models
- Generalized linear models
- Gaussian process regression

Major references:

- Chapters 14, 15, 16, 21 of [7]. Chapter 9, 11 of [9].
- Lecture notes ‘RegressionModels.pdf’.
- *R package rstanarm*: [5]
- *Gaussian process materials*: <http://www.gaussianprocess.org/>
- Exercise: Homework 4.

9 Model Checking and Comparison

- Model checking
- Model comparison

Major references:

- Chapters 6, 7 of [7]. Chapter 4 of [9].
- Lecture notes ‘ModelComparison.pdf’.

10 Mixture Models

- Finite mixture models
 - Gaussian mixture models
- Dirichlet process: [16]

– Tutorials: <http://stat.columbia.edu/~porbanz/npb-tutorial.html>

Major references:

- Chapters 22, 23 of [7].
- Lecture notes ‘MixtureModels.pdf’.

11 Missing Data Problems

- Missing data imputation [11].
- Causal inference [10, 4].

Major references:

- Chapters 18 of [7].
- Lecture notes ‘MissingData.pdf’.
- Exercise: Homework 5.

12 Approximate Inference

- Approximate Bayesian computation: [2, 13, 3]
- Variational Bayes: [1, 17]

Major references:

- Chapters 13 of [7].
- Lecture notes ‘ApproximateInference.pdf’.

References

- [1] David M. Blei, Alp Kucukelbir, and Jon D. McAuliffe. Variational inference: A review for statisticians. *Journal of the American Statistical Association*, (just-accepted), 2017.
- [2] Katalin Csilléry, Michael G. B. Blum, Oscar E. Gaggiotti, and Olivier François. Approximate Bayesian computation (ABC) in practice. *Trends in ecology & evolution*, 25(7):410–418, 2010.
- [3] Pierre Del Moral, Arnaud Doucet, and Ajay Jasra. An adaptive sequential Monte Carlo method for approximate Bayesian computation. *Statistics and Computing*, 22(5):1009–1020, 2012.
- [4] Paolo Frumento, Fabrizia Mealli, Barbara Pacini, and Donald B Rubin. Evaluating the effect of training on wages in the presence of noncompliance, nonemployment, and missing outcome data. *Journal of the American Statistical Association*, 107(498):450–466, 2012.
- [5] Jonah Gabry and Ben Goodrich. rstanarm: Bayesian applied regression modeling via stan. *R package version*, 2(1), 2016.

- [6] A. Gelman. Rstan: the R interface to stan. 2014.
- [7] Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. *Bayesian Data Analysis, Third Edition*. Chapman and Hall/CRC, 2013.
- [8] Jiqiang Guo, D. Lee, K. Sakrejda, J. Gabry, B. Goodrich, J. De Guzman, E. Niebler, T. Heller, and J. Fletcher. rstan: R interface to stan. *R package version*, 2:0–3, 2016.
- [9] Peter D. Hoff. *A first course in Bayesian statistical methods*. Springer Science & Business Media, 2009.
- [10] Roderick J Little and Donald B Rubin. Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. *Annual review of public health*, 21(1):121–145, 2000.
- [11] Roderick J. A. Little and Donald B. Rubin. *Statistical analysis with missing data*. John Wiley & Sons, 2014.
- [12] Jun S. Liu. *Monte Carlo Strategies in Scientific Computing*. Springer-Verlag New York, Inc., 2001.
- [13] Jean-Michel Marin, Pierre Pudlo, Christian P Robert, and Robin J Ryder. Approximate Bayesian computational methods. *Statistics and Computing*, pages 1–14, 2012.
- [14] Richard McElreath. Markov chains: Why walk when you can flow? 2017. <http://elevanth.org/blog/2017/11/28/build-a-better-markov-chain/>.
- [15] Walter W. Piegorsch and A. John Bailer. Appendix A: Review of probability and statistical inference. *Analyzing Environmental Data*, pages 411–439. <http://onlinelibrary.wiley.com/store/10.1002/0470012234.app1/asset/app1.pdf?v=1&t=jbwhfgrf&s=065303c8e01b85bec2be9736366d8d2f8bbdf98f>.
- [16] Yee Whye Teh. Dirichlet process. In *Encyclopedia of machine learning*, pages 280–287. Springer, 2011.
- [17] Martin J. Wainwright, Michael I. Jordan, et al. Graphical models, exponential families, and variational inference. *Foundations and Trends in Machine Learning*, 1(1–2):1–305, 2008.