Faculty of Economics
University of Tor Vergata
Rome
Bayesian Methods in Microeconometrics
May $15 - 17^{th}$ 2019

Lecturer: M. J. Weeks

Assistant Professor of Economics, University of Cambridge

This course will provide an introduction to simulation-based methods that are commonly used in microeconometrics. The emphasis will be Bayesian, although we will also contrast posterior analysis with maximum likelihood estimation.

The course covers a mix of theory and applications. We will start with a number of theoretical issues including exchangeability, prior-posterior analysis and hypothesis testing. We will also examine the fundamental problem of prior elicitation.

We present the basic ideas behind Monte Carlo integration and demonstrate the use of these techniques in both maximising a likelihood and sampling from a posterior distribution. We will examine the fundamentals of MCMC from both a theoretical and applied setting. We start with some theory, focussing on the integral transform theorem and demonstrating how this provides a route to simulating random variables. We provide a brief introduction to simple Accept-Reject sampling methods, moving onto the Gibbs sampler and Metropolis-Hastings.

Applications will be taken from discrete choice, treatment models, endogeneity and consumer heterogeneity in marketing applications. In each case we consider the particular advantages of a Bayesian approach, how to implement, and how data augmentation can be a useful tool in designing a posterior simulator.

We will also provide a short introduction to machine learning with a focus on decision trees for prediction and causal effects.

Given time constraints, we will provide examples using `Python` and `R`.

Topics: The course will select from the following topics:

1. Bayesian Methods in Microeconometrics
2. Prior Elicitation for Bayesian Econometrics
3. An Introduction to Monte Carlo Sampling
4. Posterior Sampling for Microeconometric Models: Data Augmentation and the EM Algorithm
5. Bayesian Models for Treatment Effects & Selection Models
6. An Introduction to Python

Textbooks: we will make use of the following texts:

Geweke, J., Koop, G., and H. Van Dijk (2011): *The Oxford Handbook of Bayesian Econometrics*. Oxford University Press

Koop, G. (2003): *Bayesian Econometrics*. Chichester, John Wiley & Sons.

Greenberg, E. (2004). *Introduction to Bayesian Econometrics*, Cambridge University Press. Chapters 1-4.

Lancaster, T. (2004). *An Introduction to Modern Bayesian Econometrics*. Blackwell.

Mariano, R., Schuermann, T., and M. Weeks (2008) *Simulation-Based Inference in Econometrics*. Cambridge University Press

Rossi, P. McCullogh and Allenby (2009) *Bayesian Statistics and Marketing.*. Wiley Series in Probability and Statistic.

Train, K. (2008) *Discrete Choice Methods with Simulation* . Cambridge University Press (second edition).

Evaluation: There will be a take home exam. The exam will be distributed after the last lecture. It will be due in the beginning of the Lent Term

Topic 1: Bayes, Inverse Probability and Conjugate Priors]

(a) Bayesian Inference and Heart Attacks

(b) From Inverse Probability to Bayesian Inference

(c) Bayesian Methods with Conjugate Priors

(d) Posterior Assessment for a Proportion

Readings**:**

Cornfield, J. (1967) *Bayes Theorem*, Review of the International Statistical Institute Vol. 35, No. 1 (1967), pp. 34-49.

Cornfield, J. (1951) *A method of estimating comparative rates from clinical data; applications to cancer of the lung, breast, and cervix.* J. National Cancer Inst 11, pp. 1269-75.

Fienberg, Stephen E. (2006). *When Did Bayesian Inference Become "Bayesian"?* Bayesian Analysis. 1 (1): pp. 1-40

Topic 2: Prior Elicitation for Bayesian Econometrics

(a) The Bayesian Paradigm

(b) Prior Uncertainty

(c) Reference and Jeffrey's Priors

(d) Conjugate Priors

(e) Hierarchical Priors

(d) Dirichlet Process Priors

Topic 3: Fundamentals of Bayesian Inference

(a) Bayes Theorem for Events and parameters

(b) De Finetti's Representation Theorem

(c) Parameter Uncertainty

(d) Model Uncertainty

(f) Multiplicity

(g) Hierarchical Models for Combining Data

(h) Model Selection and Model Averaging

Readings:

Koop (2008), Chapter 1.1; Lancaster (2004), Chapter 1.

Rossi et al (2009); Chapter 2.0-2.5

Practical Session 1: Priors, Posteriors and Bayes

(a) An Introduction to Python and the Python Environments

(b) Bayesian And Frequentists Notions of Probability

(c) Generating Random Variables

(d) Conjugate Model: The Beta-Binomial Model

(e) Markov Chain Monte Carlo (non-conjugate priors)

(f) Parameter Uncertainty: Bayesian and Classical Interval Estimators

Readings:

Lancaster (2004), Chapter 4

Koop (2003), Chapters 2, 5, and 12

Mariano, Schuermann, and Weeks (2008), Chapter 1;

Topic 4: An Introduction to Monte Carlo Sampling

(a) The Probability Integral Transform

(b) Rejection Sampling

(c) Variants of the Beta distribution

(d) Posterior Assessment for a Proportion

(e) Markov Chain Monte Carlo

- The Gibbs Sampler
- The Metropolis-Hastings sampler

Readings**:**

Mariano, B., Schuermann, T., and M. Weeks (2008), Chapter 1;

Geweke, J., Koop, G., and H. Van Dijk (2011): Chapter 5;

Gilks, W., Richardson, S., and D. Spiegelhalter (1996):

1. Chapter 1: Introducing Markov Chain Monte Carlo
2. Chapter 3: Markov Chain Concepts Related to Sampling Algorithms
3. Chapter 5: Full Conditional Distributions, 3, and 5

Chib, S. (1995) *Understanding the Metropolis-Hastings Algorithm and the Poor Man's Data Augmentation Algorithm*, The American Statistician Vol. 49, No. 411 (Nov., 1995).

Wei, G., and M. Tanner (1990) A Monte Carlo Implementation of the EM Algorithm and the Poor Man's Data Augmentation Algorithm, *Journal of the American Statistical Association* Vol. 85, No. 411

Practical Session 3: An Introduction to Monte Carlo Sampling

(a) Inverse CDF Method for Discrete and Continuous Variables
(b) Rejection Method
(c) Introducing Monte Carlo Sampling
(d) Inference for a mean parameter

Practical Session 3a: Markov Chain Monte Carlo: Theory

(a) Overview
(b) Building Models with `PyMC`
(c) Fitting Models with `PyMC`
(d) Markov chain Monte Carlo: the `MCMC` class
(e) Step Methods: Gibbs and Metropolis Hastings
(f) Some Examples

Readings**:**

Rossi et al (2009); Chapter 3

Lancaster (2004), Chapter 4

Practical Session 3b: Markov Chain Monte Carlo: Practice

(a) Regression
(b) Lasso Regression
(c) IV Regression
(d) Model Selection and Bayes Factors

Readings**:**

Rossi et al (2009); Chapter 6.5

Lancaster (2004); Chapter 8

Kass R, Raftery A. 1995. Bayes Factors. *JASA*

Topic 5: The EM Algorithm and Data Augmentationfor Latent Variable Models

(a) The EM Algorithm

(b) Data Augmentation for Missing Data Models

(c) Bayesian Inference for Binary Choice

(d) Bayesian Inference for the Mixed Logit Model

(e) Semiparametric Estimators and the EM Algorithm

Readings:

Geweke, J., Koop, G., and H. Van Dijk (2011): Chapter 5 (4.1)

Topic 6: Dirichlet Processes for Nonparametric Bayesian Models

(a) Heterogeneity in Econometric Models

(b) Finite Mixture Models

(c) Bayesian Histograms

(d) Dirichlet distributions, Dirichlet Processes, Dirichlet Process priors

(e) Generative distributions for Dirichlet Processes: Chinese Restaurant and Stick Breaking

(f) Dirichlet Processes and Infinite Mixture Models

(g) Semiparametric Bayesian Generealised Least Squares

Readings:

Geweke, J., Koop, G., and H. Van Dijk (2011): Chapter 6

Albert, J., and S. Chib (1993)

Koop (2003), Chapter 9.

Cameron and Trivedi (2005). Chapter 23.

Train, K. (2008).

Topic 7a: Machine Learning and Decision Trees

(a) What is Machine Learning
(b) Contrast with Traditional Econometrics
(c) What Economists Do
(d) Prediction in a Stable Environment
(e) Key Lessons for Econometrics

Topic 7b: Machine Learning and Decision Trees

(a) Machine Learning: Terminology and Concepts
(b) An Overview of Regression Trees
(c) Training, Testing and Cross Validation
(d) Adaptive versus Honest Estimation
(e) Variance Reduction Methods
(f) Variable Importance
(g) Causal Trees and Forests
(h) Application: Time of Use Tariffs and Smart Meter Data

END

Textbooks

[1] Koop, G. (2003): *Bayesian Econometrics*

[2] Rossi, P. McCullogh and Allenby (2009) *Bayesian Statistics and Marketing* , Wiley Series in probability and Statistics.

[3] Rossi, P., Robert E. McCulloch, and Greg M. Allenby. 2005. *Hierarchical Bayes: A Practitioners Guide.* mimeo, Ohio State University.

[4] Lancaster, T. (2004): *An Introduction to Modern Bayesian Econometrics*, Blackwell Publishing

[5] Gilks, W., Richardson, S., and D. Spiegelhalter (1996) *Markov Chain Monte Carlo in Practice*, Chapman and Hall

[6] Mariano, R., Schuermann, T., and M. Weeks (2008): *Simulation-based Inference in Econometrics Methods and Applications*, Cambridge University Press.

[7] Greenberg, E. (2013): *Introduction to Bayesian Econometrics*, Second Edition, Cambridge University Press.

[8] Koop, G. and Poirier, D. J. and Tobias, L. (2007) Bayesian Econometric Methods (Econometric Exercises), Cambridge University Press.

[9] Lancaster, T. (2004). *An Introduction to Modern Bayesian Econometrics.* Blackwell.

[10] Train, K. (2008) *Discrete Choice Methods with Simulation*, Cambridge University Press (second edition).

[11] Van Dijk, H.K., A. Monfort and B.W. Brown (eds) (1995). *Econometric Inference Using Simulation Techniques*, John Wiley and Sons, Chichester, West Sussex, England.

Markov Chain Monte Carlo

[1] Albert, J., and S. Chib (1993) Bayesian Analysis of Binary and Polychotomous Response Data, *Journal of the American Statistical Association*, Vol. 88, No. 422 pp. 669-679.

[2] Casella, G., George, E.I. (1992). Explaining the Gibbs Sampler. *The American Statistician*, 46, 167 174.

[3] Gelfand, A. E., and F. M. Adrian Smith (1990) Sampling-Based Approaches to Calculating Marginal Densities, *Journal of the American Statistical Association*, Vol. 85, No. 410. , pp. 398-409.

[4] Wei, G., and M. Tanner (1990) A Monte Carlo Implementation of the EM Algorithm and the Poor Man's Data Augmentation Algorithm, *Journal of the American Statistical Association* Vol. 85, No. 411 (Sep., 1990), pp. 699-704.

[5] Chib, S. (1995) *Understanding the Metropolis-Hastings Algorithm and the Poor Man's Data Augmentation Algorithm*, The American Statistician Vol. 49, No. 411 (Nov., 1995).

[6] Metropolis, N. & S. Ulam (1949). The Monte-Carlo method. *Journal of the American Statistical Association* 44:335-341.

[7] Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller & E. Teller (1953). Equation of state calculations by fast computing machines. *Journal of Chemical Physics* 21: 1087-1092.

[8] Tanner, M., and W. Wong (2010) From EM to Data Augmentation: The Emergence of MCMC Bayesian Computation in the 1980s. *Statistical Science* Vol. 25, No. 4, 506516.

Basics of Bayesian Inference

[1] Akaike H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle. In *Second International Symposium on Information Theory*, Petrov B, Csake F. (eds). Akademiai Kiado: Budapest.

[2] Geweke J. (1993). "Bayesian Treatment of the Independent Student-t Linear Model. *Journal of Applied Econometrics* **8**: S19-S40.

[3] Kass R., Raftery A. (1995). Bayes Factors. *Journal of the American Statistical Association* **90(430)**: 773-95.

[4] Leamer E.E. (1973). Multicollinearity: A Bayesian Interpretation. *Review of Economics and Statistics* **55(3)**: 371-80.

[5] Smith, A., & A. Gelfand (1992). Bayesian Statistics without Tears: A SamplingResampling Perspective. *The American Statistician*, vol. 46, Issue 2.

Selection and Treatment Effects

[1] Chib, S. and B.H. Hamilton, (2000), Bayesian analysis of cross-section and clustered data treatment models, *Journal of Econometrics*, 97(1), 25-50

[2] Munkin, M.K. and P.K. Trivedi, (2003), Bayesian analysis of a self-selection model with multiple outcomes using simulation-based estimation: An application to the demand for healthcare, *Journal of Econometrics*, 114(2), 197-220

[3] Li, M, and J. Tobias, (200*), Bayesian analysis of Treatment Effects in an Ordered Potential Outcomes Mode1, Advances in Econometris,

Semiparametric Bayesian Models

[1] Pulak Ghosh, Paramjit S. Gill, Muthukumarana, S and T. B. Swartz A Semiparametric Bayesian Approach to Network Modelling using Dirichlet Process Priors, mimeo, Emory University

[2] Conley, T., Hansen, C., McCulloch, R. & P. Rossi (2008): A semi- parametric Bayesian approach to the instrumental variable problem, Journal of Econometrics, Vol 44(1), pp. 276−305.

[3] Bajari, P, J.T. Fox, and S. P. Ryan. (2007) Linear Regression Estimation of Discrete Choice Models with Nonparametric Distributions of Random Coefficients, *American Economic Review*, 97(2): 459−463.

Models of Heterogeneity

[1] Huber, J., & Train, K. (2001). On the similiarity of classical and Bayesian estimates of individual mean partworths. *Marketing Letters*, 12, pp 259-269.

[2] Rossi, P. E., Gilula, Z., & Allenby, G. M. (2001). Overcoming scale usage heterogeneity: A Bayesian hierarchical approach. *Journal of the American Statistical Association*, 96, pp20- 31.

[3] Rossi, P. E., & Allenby, G. M. (1993). A Bayesian approach to estimating household parameters. *Journal of the Marketing Research*, 30, pp 171-182.

[4] Rossi, P. E., & Allenby, G. M. (1998). Marketing models of consumer heterogeneity *Journal of Econometrics*, vol. 89, issue 1-2, pp 57-78.

Bayesian Discrete Choice Models

[*] *CHAPTER 13*, Bayesian Methods in Cameron, C.A. and P.K. Trivedi (2005) *Microeconometrics: Methods and Applications*. Cambridge University Press.

[1] Albert, J. and S. Chib (1993), "Bayesian analysis of binary and polychotomous response data", *Journal of the American Statistical Association* 88, 669-679.

[2] Allenby, G.M. and P.E. Rossi (1999) "Marketing Models of Consumer Heterogeneity" *Journal of Econometrics* 89, 57-78.

[3] Chib, S. and E. Greenberg (1998), "Analysis of multivariate probit models", *Biometrika* 85, 347-361.

[4] Kosuke Imaia, D. van Dyk, A Bayesian analysis of the multinomial probit model using marginal data augmentation

[5] Nobile, A., (1998). A hybrid Markov chain for the Bayesian analysis of the multinomial probit model. *Statistics and Computing* 8, 229242.

[6] Nobile, A., (2000). Comment: Bayesian multinomial probit models with normalization constraint. *Journal of Econometrics* 99, 335345

[7] McCulloch, R. E., & Rossi, P. E. (1994). An exact likelihood analysis of the multinomial probit model. *Journal of Econometrics*, 64, 217-228

[8] McCulloch, R.E., Polson, N.G., Rossi, P.E., (2000). A Bayesian analysis of the multinomial probit model with fully identified parameters, *Journal of Econometrics* 99, 173-193.

[9] McCulloch, R. and P. Rossi (2000), "Bayesian analysis of the multinomial probit model" in R. Mariano, T. Schuermann and M. Weeks (eds). *Simulation-Based Inference in Econometrics*, Cambridge University Press, New York.

[10] Poirier, D.J. (1996) "A Bayesian Analysis of Nested Logit Models" *Journal of Econometrics* 75, 163-181.

Machine Learning

[1] A. Cameron, P. Trivedi. *Microeconometrics* Cambridge University Press, 2005.

[2] L. Breiman, J. Freidman, R. Olshen, C. Stone. *Classification and Regression Trees*. Klein-Verlag, 1990.

[3] *Random Forests*. `https://en.wikipedia.org/wiki/Random_forest`.

[4] *Training, Validation, and Test sets*. `https://en.wikipedia.org/wiki/Training,_validation,_and_test_sets`

[5] J. Freidman, T. Hastie, R. Tibshirani. *The Elements of Statistical Learning*. Springer, 2009.

[6] G. James, D. Witten, T. Hastie, R. Tibshirani. *An Introduction to Statistical Learning with Applications in R*. Springer, 2013.

[7] S. Athey, G. Imbens. Recursive partitioning for heterogeneous causal effects *Proceedings of the National Academy of Sciences*, 113(27):7353–7360, 2016.

[8] E. O'Neill, M. Weeks. Causal Tree Estimation of Heterogeneous Household Response to Time-Of-Use Electricity Pricing Schemes arXiv:1810.09179v1, 2018.

[9] S. Athey, G. Imbens, Y. Kong, V. Ramachandra. An Introduction to Recursive Partitioning for Heterogeneous Causal Effects Estimation Using causalTree package. https://github.com/susanathey/causalTree/blob/master/briefintro.pdf, 2016.

[10] S. Athey, S. Wager. Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 2017.

[11] Y. Lin, J. Yongho Random forests and adaptive nearest neighbors *Technical Report No. 1055. University of Wisconsin, 2002* http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.153.9168

[12] S. Russell, P. Norvig Artificial Intelligence: A Modern Approach *3rd edition, 2009*

LASSO [NOT COVERED in 2019]

[1] Efron, B., Hastie, T., Johnstone, I. and R. Tibshirani, (2004). Least angle regression (with discussion), *Annals of Statistics* 32(2): 407- 499

[2] Friedman, Jerome; Hastie, Trevor; Tibshirani, Robert (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Second Edition (Springer Series in Statistics) (Kindle Locations 13024-13026). Springer - A. Kindle Edition.

[3] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *J. Royal. Statist. Soc B.*, Vol. 58, No. 1, pages 267-288).

[4] Hofmarcher, P., Cuaresma, J., Grun, B., and K. Hornik (2015), Last Night a Shrinkage Saved My Life: Economic Growth, Model Uncertainty and Correlated Regressors, *Journal of Forecasting*, Vol. 34, pages 133144

[5] Park, T., and Casella, G. (2008). The Bayesian lasso. *Journal of the American Statistical Association*, 103(482), 681-686.

[6] O'Hara, R. B., & Sillanpaa, M. J. (2009). A review of Bayesian variable selection methods: what, how and which. *Bayesian analysis*, 4(1), 85-117.

[7] Kyung, M., Gill, J., Ghosh, M., & Casella, G. (2010). Penalized regression, standard errors, and Bayesian lassos. *Bayesian Analysis*, 5(2), 369-411.

[8] Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(3), 273-282.