Fall 2020

Stony Brook University

Department of Ecology & Evolution

College of Arts and Sciences

Course Title: Bayesian Data Analysis and Computation

Course Instructor: Dr. Heather Lynch

Instructor contact information:

Dr. Heather Lynch

Email: heather.lynch@stonybrook.edu

Work Phone: 2-9508

Office Hours Location: Zoom [Link TBD]

Office Hours: Friday 3:30-5:00 pm (subject to change depending on travel and unavoidable scheduling conflicts)

TA contact information:

Bilgecan Sen

Email: bilgecan.sen@stonybrook.edu

Office Hours Location: Zoom [Link TBD]

Office Hours: Thursday 9:00-10:30 am

Important Note: Every effort will be made to avoid changing the course schedule, but the possibility exists that unforeseen events will make syllabus changes necessary. It is your responsibility to check the course Slack page for corrections or updates to the syllabus. Any changes will be clearly noted in course announcements or through Stony Brook email.

COURSE DESCRIPTION:

An applied course in Bayesian data analysis and hierarchical modeling for advanced graduate students in Ecology& Evolution or related sciences. Topics will include probability theory, Bayesian analysis, and MCMC methods such as Gibbs sampling and Metropolis-Hastings sampling, as well as applied issues regarding the choice of prior distributions, posterior convergence, censored and missing data, and model checking and comparison. The course will be taught using JAGS as accessed via the R package R2jags.

TECHNICAL REQUIREMENTS:

This course will occur synchronously on Zoom. You will need a computer with a camera and microphone and an internet connection. You will also need to have R, Excel, and JAGS installed on your computer. While the course will be recorded, class participation is an essential element of the course and all students are expected to attend synchronously; accommodations for absences will be considered on a case-by-case basis.

*Course Pre/co-requisites*

BEE 552 (Biometry) or equivalent. The course assumes a solid understanding of calculus and R proficiency.

COURSE OBJECTIVES:

This is an advanced graduate course focusing on the theory and application of Bayesian statistical methods. At the end of this course, students will:

1. know Bayes theorem and the conceptual differences between frequentist and Bayesian approaches (Weeks 1&2)

2. understand prior distributions and conjugate priors (Week 3)

3. know how to construct prior distributions based on expert opinion (Week 3)

4. know how to interpret and summarize the posterior distribution (Week 3&9)

5. understand the basic principles and be able to write R code for a variety of MCMC sampling techniques, including Monte Carlo integration, importance sampling, Gibbs sampling, Metropolis-Hastings sampling (Weeks 4&6)

6. know how to write code in JAGS to do Bayesian analyses (Week 7)

7. do standard statistical analyses using a Bayesian approach (e.g., linear regression, t-tests, mixed-models) (Weeks 8&15)

8. test for convergence of the posterior distribution (Week 9)

9. know how to do statistical inference using posterior distributions and Bayes factors (Week 10)

10. know how to compare model fit using DIC and other Information Theoretic approaches (Week 11)

11. be able to construct hierarchical Bayesian models (Weeks 12&13)

12. know how to deal with common challenges of ecological modeling such as censored and missing data (Week 14-15)

*Attendance and Make Up Policy*

Late work will not be accepted. Attendance at all exams is mandatory. In the case of 1) verifiable illness, 2) verifiable family emergency, 3) University-sanctioned religious holiday, or 4) participation in official University-sponsored events (for documented student athletes only), excuse must be documented on official letterhead (as appropriate) and will be verified by the instructor.

*Description and schedule of Required Readings and/or Assignments.*

Readings: Readings will be assigned primarily from

McCarthy, M.A. 2007. Bayesian Methods for Ecology. Cambridge University Press, Cambridge, UK.

Hobbs, N.T., and M.B. Hooten. Bayesian Models. Princeton University Press, Princeton, NJ.

Additional readings from the primary literature will occasionally be assigned for discussion in recitation as described in the tentative syllabus below.

**Tentative Syllabus**:

|  |  |  |  |
| --- | --- | --- | --- |
| **Week** | **Dates** | **Topic** | **Text Readings**  Papers in [] are suggested but not required; papers with a \* involve advanced material that may take several weeks to sink in |
| **Week 1** | Aug 24  Aug 26 | Introduction to Bayesian Statistics | McCarthy 1-2  \*Christensen (2005)  Clark and Gelfand (2006)  [Flam NYTimes article]  Stephens et al. (2007) |
| **Week 2** | Aug 31  Sept 2 | Bayesian inference and Prior distributions | McCarthy 3  [Fenton and Neil (2012)]  Ellison (1996)  \*Ellison (2004)  H&H Chapter 5  Berger and Berry (1988) [a re-read from Biometry]  McCarthy Appendix B  Lunn et al. 5 |
| **Week 3** | ~~Sept 7~~  Sept 9 | Prior distribution lab | Kuhnert et al. (2010)  \*Lambert et al. 2005  \*Senn 2007  \*Lambert et al. 2008 |
| **Week 4** | Sept 14  Sept 16 | Introduction to Monte Carlo: Rejection Sampling, Monte Carlo Integration and Importance Sampling | Lunn et al. 1.4  MCMC-UseR pg. 51-79  Smith and Gelfand (1992)  H&H Chapter 7 |
| **Week 5** | Sept 21  Sept 23 | MCMC methods: Gibbs sampling and Metropolis-Hastings | McCarthy Appendix C  Lunn et al. 4  Clark Lab 5  Kery 2  MCMC-UseR pg. 126-167 |
| **Week 6** | Sept 28  Sept 30 | Midterm & Introduction to JAGS | McCarthy Appendix A  JAGS Primer |
| **Week 7** | Oct 5  Oct 7 | Bayesian regression models | McCarthy 5 (all except pp. 130-148)  Lunn et al. 6.1-6.4  Kery 8  H&H Chapter 8 |
| **Week 8** | Oct 12  Oct 14 | Bayesian GLMs and mixed-effects models | McCarthy 5 (pp. 130-148)  Kery 13-14, 16 |
| **Week 9** | Oct 19  Oct 21 | Hierarchical modeling and spatial/phylogenetic models | Gelman and Hill Chapter 11 & 16  Latimer et al. (2006)  Nice et al. (2014)  H&H Chapter 6 |
| **Week 10** | Oct 26  Oct 28 | Practical issues in Bayesian computation (convergence diagnostics, imputation of missing data) | Lunn et al. 4.3-4.5  MCMC-UseR Chapter 8 (starts Pg. 210)  Brooks et al. 1.11 (pp. 17-21)  Nakagawa and Freckleton (2008)  Gelman and Hill Chapter 25 |
| **Week 11** | Nov 2  Nov 4 | Population dynamics/Time series/State-space models | McCarthy 9  Bolker Chapter 11  Clark and Bjørnstad (2004)  For problem set: Clark et al. (2010) |
| **Week 12** | Nov 9  Nov 11 | Mark-recapture and occupancy modelling | McCarthy 7  Gimenez et al. (2009)  McCarthy and Masters (2005)  Rotella Notes on CJS  Royle and Dorazio Ch. 3&9  Royle and Kery (2007) |
| **Week 13** | Nov 16  Nov 18 | Approximate Bayesian Computation | Beaumont (2010)  Beaumont and Rannalla (2004) |
| **Week 14** | Nov 30  Dec 2 | Hypothesis testing & model evaluation | McCarthy 4 (pp. 94-108)  Ward (2008)  Gelman and Shalizi (2013)  Spiegelhalter et al. (2014)  Lunn et al. 8.1-8.6  Greenland and Poole (2013)  Pitt and Myung (2002) |
| **Week 15** | Dec 7 | Bayesian Multimodel Inference | McCarthy (pp. 108-118)  Lunn et al. 8.7-8.10  Link and Barker (2006)  Link and Barker Chapter 7  Tenan et al. (2014)  H&H Chapter 9 |

Partial list of additional readings:

Beaumont, M. A. 2010. Approximate Bayesian Computation in Evolution and Ecology. Annual Review of Ecology, Evolution, and Systematics 41: 379-406.

Beaumont, M. A., and B. Rannala. 2004. The Bayesian revolution in genetics. Nature Reviews Genetics 5: 251-261.

Clark, J.S., and A.E. Gelfand. 2006. A future for models and data in environmental science. TREE 21(7): 375-380.

Ellison, A. M. 1996. An introduction to Bayesian inference for ecological research and environmental decision making. Ecological Applications 6(4): 1036-1046.

Ellison, A. M. 2004. Bayesian inference in ecology. Ecology Letters 7: 509-520.

Gelman, A., and C.R. Shalizi. 2013. Philosophy and the practice of Bayesian statistics. British Journal of Mathematical and Statistical Psychology 66: 8-38.

Gimenez, O. et al. 2009. WinBUGS for population ecologists: Bayesian modeling using Markov Chain Monte Carlo methods. In: D.L. Thompson et al. (Eds.) Modeling Demographic Processes in Marked Populations. Environmental and Ecological Statistics 3: 883-915.

Hoeting, J.A., D. Madigan, A.E. Raftery, and C.T. Volinsky. 1999. Bayesian model averaging: A tutorial. Statistical Science 14(4): 382-417.

Kass, R. E. , and A. E. Raftery. 1995. Bayes Factors. Journal of the American Statistical Association 90(430): 773-795.

Kuhnert, P.M., T.G. Martin, and S.P. Griffiths. 2010. A guide to eliciting and using expert knowledge in Bayesian ecological models. Ecology letters 13:900-914.

Link, W.A., and R. J. Barker. 2006. Model weights and the foundations of multimodel inference. Ecology 87(10): 2626-2635.

Smith, A.F.M., and A.E. Gelfand. 1992. Bayesian statistics without tears: A sampling-resampling perspective. The American Statistician 46(2): 84-88.

Ward, E.J. 2008. A review and comparison of four commonly used Bayesian and maximum likelihood model selection tools. Ecological Modelling 211: 1-10.

Assignments:

There will be weekly problem sets involving Bayesian theory and computation.

*Exams*

The midterm exam will be based on all information presented up through Week 6 and will be designed to take 2 hours. The final exam will include all material presented in lecture and will be designed to take 3 hours.

GRADING:

The course grade will be based on the following components:

|  |  |
| --- | --- |
| Item | Percent |
| Problem Sets (3% each x 12 problem sets) | 36 |
| Midterm exam | 20 |
| Project proposal | 6 |
| Final project | 28 |
| Final project peer review | 10 |

Grades are based on the following scale:

A = 930-1000, A- = 900-929

B+ = 880-899, B = 830-879, B- = 800-829

C+ = 780-799, C = 730-779, C- = 700-729

D+ = 680-699, D = 630-679, F <630

MEETING SCHEDULE

Mid-term exam: TBD

Final exam: TBD

CLASS PROTOCOL

All electronic devices are to be turned off during class unless advance permission is given by the instructor.

CLASS RESOURCES

Blackboard will be used as the primary means of distribution for readings from the primary literature and submission of assignments.

If you have a physical, psychological, medical or learning disability that may impact your course work, please contact Disability Support Services, ECC (Educational Communications Center) Building, room128, (631) 632-6748. They will determine with you what accommodations, if any, are necessary and appropriate. All information and documentation is confidential.

Students who require assistance during emergency evacuation are encouraged to discuss their needs with their professors and Disability Support Services. For procedures and information go to the following website: http://www.stonybrook.edu/ehs/fire/disabilities ]

Each student must pursue his or her academic goals honestly and be personally accountable for all submitted work. Representing another person's work as your own is always wrong. Faculty are required to report any suspected instances of academic dishonesty to the Academic Judiciary. Faculty in the Health Sciences Center (School of Health Technology & Management, Nursing, Social Welfare, Dental Medicine) and School of Medicine are required to follow their school-specific procedures. For more comprehensive information on academic integrity, including categories of academic dishonesty, please refer to the academic judiciary website at <http://www.stonybrook.edu/uaa/academicjudiciary/>

Stony Brook University expects students to respect the rights, privileges, and property of other people. Faculty are required to report to the Office of Judicial Affairs any disruptive behavior that interrupts their ability to teach, compromises the safety of the learning environment, or inhibits students' ability to learn. Faculty in the HSC Schools and the School of Medicine are required to follow their school-specific procedures.