

# Syllabus for STA360/601

## Bayesian Inference and Modern Statistical Methods

### Spring 2015, Duke University

## 1 General information

Lectures: Tuesdays and Thursdays, 4:40–5:55 PM, Old Chem 116

Labs:

360-01L: Thursday 11:45–1:00, Old Chem 101

601-01L: Thursday 1:25–2:40, Old Chem 101

601-02L: Thursday 3:05–4:20, Old Chem 101

Course website: <https://sakai.duke.edu>

Textbook: *A First Course in Bayesian Statistical Methods*, Peter D. Hoff, 2009, New York: Springer. (*Note: We will only loosely follow the book.*)

Optional supplementary text: *Bayesian Data Analysis*. Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., & Rubin, D.B. (2013). CRC press.

### Primary instructor

Jeff Miller

Office: Old Chem 216

Office hours: 3:30-4:30 PM Tuesdays, 6:00-7:00 PM Thursdays

Department: Statistical Science

Email address: [jeff.miller@duke.edu](mailto:jeff.miller@duke.edu)

Website: <https://stat.duke.edu/~jwm40/>

### Teaching assistants

Christoph Hellmayr

Email address: [christoph.hellmayr@duke.edu](mailto:christoph.hellmayr@duke.edu)

Ksenia Kyzyurova

Email address: [ksenia@stat.duke.edu](mailto:ksenia@stat.duke.edu)

Michael Lindon

Email address: [m33@stat.duke.edu](mailto:m33@stat.duke.edu)

## 2 Schedule / Important dates

### Key dates:

Midterm exam: Tuesday, Feb 17, usual class time and location

Final exam: Friday, May 1, 7:00–10:00 PM, Old Chem 116

### Class will be held on the following dates:

Jan 8, 13, 15, 20, 22, 27, 29

Feb 3, 5, 10, 12, 17 (midterm), 19, 24, 26

Mar 3, 5, 17, 19, 24, 26, 31

Apr 2, 7, 9, 14

Apr 16 and 21: Optional review session, at the usual class time and location.

## 3 Grades

Your overall score for the course will be determined by:

25% Homework

25% Labs

20% Midterm

25% Final

5% Class participation.

An overall score of  $s$  will result in a grade of:

A if  $90 \leq s \leq 100$

B if  $75 \leq s < 90$

C if  $60 \leq s < 75$

D if  $50 \leq s < 60$

F if  $0 \leq s < 50$

or, for those taking the course on a Satisfactory/Unsatisfactory basis:

S if  $60 \leq s \leq 100$

U if  $0 \leq s < 60$ .

For graduate students, it appears that there is no “D” grade (only A, B, C, or F)—consequently, in this case anything between 0 and 60 is an F.

Note: If for some reason the scores are all significantly lower than expected, these ranges will be adjusted to something reasonable.

### Midterm grades

After the midterm exam, you will be given a midterm grade assessing your overall performance so far. If you are an undergraduate, this will also be sent to the registrar. This does not go on your transcript. The main purpose of this is to let you know how you are doing in the class.

## Cumulative final

The final exam will cover material from the whole semester.

## Policy on grade changes after the end of the semester

From the faculty handbook:

*It is important to note that with the exception of I [Incomplete Work] grades and X [Absence from Final Examination] grades, changes in grades may be made by the instructor only because of an error in calculation or an error in transcription. Changes in grades may not be based on the late submission of required work, the resubmission of work previously judged unsatisfactory, or on additional work. No changes may be made in a grade after the end of the semester following the one for which the grade was assigned, although cases of error discovered after the deadline may be appealed by the student or the instructor to the Office of the Provost.*

## 4 Homework and labs

Homework will be assigned regularly, and there will typically be a weekly lab assignment. Submit your homework and lab assignments electronically via the course website. Your solutions to mathematical exercises can be typed or handwritten, but must be clear and legible, otherwise no credit can be given. To electronify handwritten solutions, there are scanners available in the library, or you can use a smart phone (there are scanner apps to handle multiple pages). For programming exercises, include your source code (typed) as well as any supporting derivations, written out separately from the code. You are free to use any programming language you choose.

### Policy on late submissions

Late submissions within 24 hours after the deadline will receive only up to 75% credit. Work submitted later than 24 hours after the deadline will receive no credit. Submissions will be timestamped, so make sure to leave enough time to submit your assignments before the deadline. If, for some reason, you experience technical difficulties posting your assignment to the website, email it to one of the TA's, and the timestamp of the email will be used to determine whether it was on-time. If you are unable to finish the assignment on-time, you should still submit whatever you have completed—partial credit is better than nothing.

### Policy on missed work due to extraordinary circumstances

In case of illness, personal emergencies, religious observation, varsity athletic participation, or other extraordinary circumstances, contact me (Jeff) as soon as possible, and if you are an undergraduate follow the guidelines here: <http://trinity.duke.edu/undergraduate/academic-policies/missing-work-classes>. If officially approved, you will have the opportunity to make up missed work due to such circumstances. Note that being busy is not an extraordinary circumstance.

## 4.1 Homework

You are free to discuss homework problems with other people, HOWEVER, when you sit down to work out and write up your solutions, you must do this by yourself, without referring to solutions (or any notes related to solutions) provided by anyone else. Each student must turn in her/his own solutions. Two or more names on one assignment is not acceptable.

## 4.2 Labs

Lab assignments will focus more on practice with hands-on implementation and programming. You are free to discuss and work together on lab assignments, however, each student must turn in her/his own work. Two or more names on one assignment is not acceptable.

There will be weekly lab sessions, in which the TA's will discuss and provide assistance with the lab assignments. Attendance at lab sessions is not mandatory, however, you are required to submit the lab assignments.

## 5 Lectures and in-class exercises

Lecture notes will be posted on the course website. Most lectures will start with a short exercise. This only counts toward class participation—if you are there and you do the exercise, you get full credit. (If you have to miss class for a good reason but you really want to make up the in-class exercise, e-mail me (Jeff) within 24 hours.)

The main point of these exercises is communication: We communicate to you what you will be expected to know, so you can gauge your level of understanding—meanwhile, you communicate to us what you are understanding and what you are struggling with.

## 6 Topics to be covered

The course is designed to provide in-depth coverage of essential core topics, as well as a high-level overview of a wide range of other topics. The following is a rough outline of what we plan to cover.

### Foundations

Bayes' theorem, Definitions & notation, Decision theory  
Beta-Bernoulli model, Gamma-Exponential model, Gamma-Poisson model

### Background and motivation

What is Bayesian inference? Why use Bayes? A brief history of statistics

### Exponential families

One-parameter exponential families, Natural/canonical form, Conjugate priors  
Multi-parameter exponential families, Motivations for using exponential families

## **Univariate normal model**

Normal with conjugate Normal-Gamma prior, Sensitivity to outliers

## **Conditional independence relationships**

Graphical models, De Finetti's theorem, exchangeability

## **Monte Carlo approximation**

Monte Carlo, rejection sampling, importance sampling

## **Gibbs sampling**

Markov chain Monte Carlo (MCMC) with Gibbs sampling, Markov chain basics, MCMC diagnostics

## **Multivariate normal model**

Normal distribution, Wishart distribution, Normal with Normal $\times$ Wishart prior

## **Linear regression**

Linear regression, basis functions, regularized least-squares, Bayesian linear regression

## **Hierarchical models and group comparisons**

Hierarchical models, comparing multiple groups

## **Bayesian hypothesis testing**

Testing hypotheses, Model selection/inference, Variable selection in linear regression

## **Priors**

Informative vs. non-informative, proper vs. improper, Jeffreys priors

## **Metropolis–Hastings MCMC**

Metropolis algorithm, Metropolis–Hastings algorithm

## **Generalized linear models (GLMs)**

GLMs and examples (logistic, probit, Poisson)

## A sampling of topics beyond the scope of this course

Overview of further interesting and current topics

## 7 Supplementary references

If you would like to have further references in addition to the course textbook (always a good idea), the following are recommended. If you only get one of these, I would recommend Gelman et al., *Bayesian Data Analysis*.

### Bayesian statistics

*Bayesian Data Analysis*. Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., & Rubin, D.B. (2013). CRC press.

*The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*. Robert, C. P. (2001). Springer Texts in Statistics.

*Statistical Decision Theory and Bayesian Analysis*. Berger, J.O. (1985). Springer.

### Statistics in general

*Statistical Inference (Vol. 2)*. Casella, G., & Berger, R.L. (2002). Pacific Grove, CA: Duxbury.

*All of Statistics: A Concise Course in Statistical Inference*. Wasserman, L. (2004). Springer.

### Probability

*Probability and Random Processes*. Stirzaker, D. & Grimmett, G. (2001). Oxford Science Publications.

### Machine learning

*Pattern Recognition and Machine Learning*. Bishop, C.M. (2006). New York: Springer.