Math for Political Science
August 14 - August 22, 2012

http://sites.duke.edu/psmathcamp/

INSTRUCTOR 1: Florian Hollenbach (fmh3@duke.edu)
INSTRUCTOR 2: Josh Cutler (josh.cutler@duke.edu)

FACULTY ADVISOR: Michael Ward **EMAIL:** mw160@duke.edu

LOCATION
(AM Sessions) Old Chem 116 (double check with orientation schedule)
(PM Sessions) Old Chem 116 (double check with orientation schedule)

COURSE STRUCTURE: The class will usually meet twice a day[1], 9:00am - 12:00pm and 2:00pm - 5:30pm. The course is not for credit. There will be no grades, but we will have a final test. Since mastering the basic concepts and skills taught in this course is essential for future coursework and graduate training, students will be expected to invest significant time and energy. The more you put into the class, the more you will get out of it! At a minimum, students at all levels will be expected to:

- do the assigned readings before the lectures;
- complete (or at least attempt to complete) assigned problem sets;
- participate in class discussions and come prepared with questions.

Morning lectures will focus on fundamental mathematical concepts that are used in statistics and formal modeling. Afternoon sessions will involve a practicum where you will work with your fellow students to develop basic computer skills and to complete problem sets. The goal of these sessions will be to gain practical experience in applying the basic skills from the morning lectures, and to preview the kinds of tasks you will be doing in future classes and in your own research.

LEARNING OBJECTIVES: The purpose of this course is to provide students with the basic mathematical and computer skills needed for the introductory statistics and formal modeling courses offered at Duke. In addition, the aim will be to provide you with a foundation for acquiring the basic mathematical literacy needed to engage the modern political science research you will be encountering in all of your courses. Each student will be coming to this class with different levels of mathematical training and skills. The purpose of this course is not to attempt to even the playing field. Rather, the course will provide all students with the opportunity to advance, no matter where they start. Students who have little (or no) training will gain the basic skills needed to take the introductory methods courses offered in the department. Students who have learned

[1]Please note deviations from this schedule in the class schedule below.

these materials before will have the chance to learn concepts at a deeper level and to better understand the most fundamental ideas used in graduate- level political science research. Advanced students may gain the confidence needed to take graduate classes in statistics and economics. We are not expecting you to *completely* master all of the materials we are covering in just six days. But we are expecting you to push yourself and take advantage of this opportunity to get a leg up in your methods training. Only a very few people in the world fully understand advanced math concepts right away. The rest of us have to work at it. In many cases, these ideas only really make sense the third or fourth time we learn them. In other words, getting a basic handle on these concepts is a matter of sustained effort, not native talent. This week is designed to give you a head start on this process. You will probably not have many opportunities to focus exclusively on these ideas on a daily basis once your classes get started, so take advantage of this opportunity by preparing before you show up in August and applying yourself during the week.

TEXTBOOKS: The required text for this course is the textbook by Jeff Gill. The course will use the *second edition* of this text. You might want to read through the chapters we will cover before you arrive in addition to reading them before each lecture.

No textbooks are required for the programming portion of the course. The R programming language has a wealth of documentation and tutorials online and an important outcome of the course will be learning where and how to find help when you need it. If you still feel that a book will help you, there are some recommendations below.

Required

- Gill, Jeff. 2006. *Essential Mathematics for Political and Social Research.* Cambridge, England: Cambridge University Press. Link

Strongly Recommended

- Kadane, Joseph B. *Principles of Uncertainty.* Project Euclid. Available online: Link

Additional References

- Adler, Joseph. 2010. *R in a Nutshell by Joseph Adler.* O'Reilley Media. Link
- Braun, John W. and Duncan J. Murdoch. 2008. *A First Course in Statistical Programming with R.* Cambridge, England: Cambridge University Press. Link
- Maindonald, John and W. John Braun. 2010. *Data Analysis and Graphics Using R.* Cambridge, England: Cambridge University Press. Link
- Simon, Carl P. and Lawrence Blume. 1994. *Mathematics for Economists.* New York, USA: W.W. Norton & Company.
- Edwards, C. Henry and David E. Penney. 2002. *Calculus.* Upper Saddle River, NJ: Prentice Hall.
- Poole, David. 2006. *Linear Algebra. A Modern Introduction.* Thomson.

- Ross, Sheldon. 2006. *A First Course in Probability.* Upper Saddle River, NJ: Prentice Hall.
- Casella, George and Roger L. Berger. 2002. *Statistical Inference.* Pacific Grove, USA: Duxbury.
- de la Fuente, Angel. 2000. *Mathematical Models for Economists.* Cambridge, England: Cambridge University Press.
- Sundaram, Rangarajan K. 1996. *A First Course in Optimization Theory.* Cambridge, England: Cambridge University Press.
- Harville, David A. 2008. *Matrix Algebra From A Statistician's Perspective.* Springer.

ADDITIONAL RESOURCES: Over the coming weeks we will be posting links to additional resources on the class website. These supplemental materials will be for both true beginners and the most advanced students. For people who have not taken math since freshman year, we will be posting online lectures, tutorials, and free books. Reviewing some of these materials in advance will help you get the most out of the class. The rest will be helpful resources for students who want to push beyond the course materials to more advanced topics. We will also post all of the prepared lecture notes as well as the data files and R scripts we will cover in class.

A NOTE ON COMPUTER: You are **not** required to have your own laptop. However, if you have one, please bring it to the afternoon lectures (and install R before the first class). On the other hand, please do not bring out laptops, phones, tablets, etc. during the morning lectures.

CLASS SCHEDULE

Tuesday, August 14

Morning Session, 9:00-noon: Introduction and the basics

- **Read: Gill Chapter 1**
- Why math?
- Arithmetic principles
- Notation
- What is a function?
- Solving an equation
- Polynomial functions
- Logarithms and exponents
- **Homework Gill Chapter 1:** 1.1, 1.2, 1.3, 1.4, 1.6, 1.8, 1.9, 1.10, 1.16, 1.17, 1.21, 1.25

Afternoon Session, 2:00-5:30pm: Introducing R

- Getting and installing **R** RStudio
- **R** as a calculator;
- Introduction to object-oriented languages;
- Workspace & object assignment;
- naming rules;
- Types of operations: maths, logical, relational;
- types of objects;
- Objects : Functions :: Nails : Hammers
- Getting help: CRAN, R-SEEK, ? and ??;
- Vector Operations: making and defining vectors, removing objects;
- element-wise operations & common vector functions (`sum()`, `mean()`, `prod()`, etc.)
- Order of operations and programming **R** to do math;
- **Homework**: Write code for "complex" math problems, ex.: $\sqrt{\frac{\cos(\frac{\pi}{2})}{e^3}}$

Wednesday, August 15

Morning Session, 9:00-noon: Basic calculus

- **Read: Gill Chapter 5, sections 5.1 – 5.4**
- Homework Problems
- Limits and sequences
- What is a derivative?

- Basic derivative rules
- L'Hospital's Rule
- **Homework Gill Chapter 5:** 5.1, 5.2, 5.3, 5.5, 5.6, 5.7, 5.8

Afternoon Session, 2:00-5:30pm: Vectors and Plots

- Vectors, continued: indexing and partitions;
- Symbolic logic required for partitions;
- cbind() and data.frames
- Summary statistics using `ChickWeight` data;
- Introduction to `plot()` - axes, scaling, labels, titles, legends, `points()`, `lines()`, `identify()`, `boxplot()`, `hist()`, `density()`.

Thursday, August 16

Morning Session, 9:00-noon: Integrals and Series

- **Read: Gill Chapter 5, , sections 5.5 – 5.7**
- Riemann sums and integrals
- The Fundamental Theorem of Calculus
- Integration rules
- **Homework Gill Chapter 5 & 6:** 5.10, 5.11, 5.13 (except for the third in the top row), 5.14 (ignore the trig functions)

Afternoon Session, 2:00-5:30pm: Sampling, Functions *I* and Data Management

- `sample()`, `rnorm()`, `runif()`,subsetting data;
- Drawing and summing random normal variables;
- plotting random variables;
- Functions: defining, purpose, writing our own functions;
- `ifelse()`,
- Data Management: mapping network directories, `setwd()`, packages, `library(foreign)`, reading in external data (read.csv, read.table, read.dta)
- Recode, reshape, generating new variables, `complete.cases()`, dealing with NAs, variable transformations, factors, sorting, local v. global objects;
- **Homework**: Summary statistics from small ANES or CCES data set.

Friday, August 17

Morning Session, 10:00-noon: Multivariate Calculus

- **Read: Gill Chapter 6, sections 6.1. – 6.6**

- Partial derivatives
- Optimization basics
- Optimization example from game theory
- Multidimensional integrals
- Finite and infinite series
- **Homework Gill Chapter 6:** 6.1, 6.7, 6.8, 6.9, 6.11 (ignore trig problems and 4th problem in left column)
  In addition:

  1. Show that for $f(x,y) = 3x^2y + 5y^4x^{\frac{2}{3}} - 2x$
     $\frac{\partial^2 f(x,y)}{\partial x \partial y} = \frac{\partial^2 f(x,y)}{\partial y \partial x}$ and $\frac{\partial^4 f(x,y)}{\partial x^2 \partial y^2} = \frac{\partial^4 f(x,y)}{\partial y^2 \partial x^2}$
  2. Maximize $f(x_1, x_2)=x_1x_2$ s.t. $h(x_1, x_2)\ x_1 + 4x_1=16$

Afternoon Session, 2:00-5:30pm: Functions *II* and Basic Programming

- Review functions;
- Branching with `if()`;
- Looping with `for()` and `while()`;
- debugging;
- uses for programming: tables, graphics, MLE, simulation;
- the `apply()` family: `tapply()`, `sapply()`, `mapply()`
- **Homework**: Elevator Simulation Problem... probably something slightly easier.

Monday, August 20

Morning Session, 9:00-noon: Probability Theory

- **Read: Gill Chapter 7, Kadane Chapter 1, sections 1.1-1.3, Chapter 2, sections 2.1, 2.4, 2.5, 2.6**
- Why do we need probability and statistics?
- Counting rules
- Sets and set operations
- What is a probability? Axioms of probability
- Conditional probability
- Bayes' rule
- Independence
- **Homework Gill Chapter 7:** 7.1, 7.5, 7.7, 7.8, 7.9, 7.13, 7.15, 7.21
  In addition: Explain in your own words what is meant by the conditional probability of A given B.

Afternoon Session, 2:00-5:30pm: LaTeX, Tables and `ggplot`

- Introduction to LaTeX;
- Tables: frequency tables and summary statistics, proportion tables, `xtable()` and `apsrtable()`;
- Lists: making lists, `which()`, `which.min()`, etc.
- Plotting II: using `ggplot`

Wednesday, August 22

Morning Session, 9:00-noon: Random variables

- **Read: Gill Chapter 8, Kadane Chapter 1, sections 1.4, Chapter 2, section 2.11**
- Making life convenient: random variables
- Levels of measurement
- Continuous and discrete RVs
- Probability mass and density function
- Cumulative distribution function
- Joint distributions, conditional distributions
- Expected value and its properties
- Higher moments of distributions
- Covariance and correlation
- Summarizing observed data
- Common distributions and their use
- **Homework Gill Chapter 7 & 8:** 7.17, 7.23, 8.1, 8.5, 8.11, 8.12
- **Homework Kadane Chapter 2, section 2.11:**Exercise 2.11.3 - third problem

Afternoon Session, 2:30-3:45pm: Putting it all together...

- Arrays and Matrices: matrix manipulation, matrix algebra, eigenvalues, determinants, etc.
- Data Management and hands on frustration: recoding, dealing with NAs... starting from a raw data set and trouble-shooting problems.
- Goals: read in data set, create new variables, generate summary statistics and run and report a bivariate regression.
- Plotting and writing up results... a one problem version of a 318 homework.