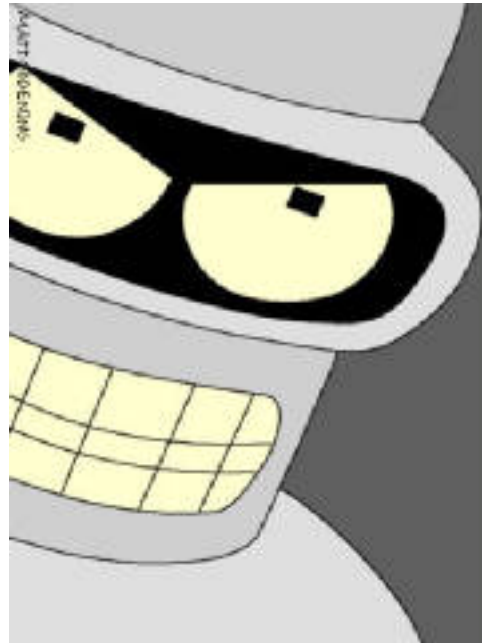


ECS 256

Norm Matloff

Fall 2010



**FOLLOW
DIRECTIONS!**

Contents

1	Things You Need to Know Right Away	3
2	Prerequisites	3
3	Required Course Materials	4
4	Role of the Textbook	4
5	Course Content	4
5.1	Topic Coverage	4
5.1.1	Review of Undergraduate Prerequisite Material	4
5.1.2	Core Course Topics	5
5.2	What You Will Get Out of the Course	5
6	Homework	5
6.1	Amount of Work	5
6.2	Writeups	6
6.3	Homework Partners	6
6.4	Submitting Homework	6
7	Quizzes and Exam	7
8	Grading	7
9	Consultation	7
9.1	Office and Office Hours	7
9.2	E-Mail	7
10	The Class Blog	8
11	Machines	8
11.1	Linux PCs in CSIF	8
12	Analytical Work Must Be Done in \LaTeX	8

1 Things You Need to Know Right Away

I teach this course (and all courses) differently from the way professors do. Please note:

- There are several Quizzes during the quarter, and a written, in-class, GROUP exam at the end of the course.
- It is REQUIRED that you read the textbook.
- Homework is done in TEAMS.
- All Tests (Exams and Quizzes) are open book.
- The textbook is open source, i.e. FREE. It is downloadable from the Web, or for sale (at printing cost) in the campus bookstore.
- You must have your own (not shared) hard copy of the entire textbook, and bring it to all Tests.

2 Prerequisites

The main required background is STA 131A (or equivalent, e.g. ECS 132 or MAT 135A), a *calculus-based* course in probability.

That means that you have at some point in the past studied (and likely forgotten, which is OK; see 5.1 below): random variables; (one-dimensional) densities; famous parametric families of distributions (at least binomial, geometric, Poisson, normal, exponential, uniform.)

This also implies that you will be able (again, after refreshing your memory) to do:



- $\int_0^\infty \lambda e^{-\lambda t} dt$

Derivatives, integrals, infinite series.

- $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} \begin{pmatrix} e \\ f \end{pmatrix}$

Basic matrix operations, i.e. addition, multiplication and inverse. (There is a review at the end of our textbook.)

It is also assumed that you have reasonable programming and debugging skill; basic awareness of the concepts of bits/bytes, memory addresses and data structures; experience in writing code to read and write files.

The  language and L^AT_EX word processing software that we will use in this class are available for every major platform—Linux, Windows and Macs. However, **I strongly prefer to that you run on Linux, as that would make it easier to help you fix your bugs.** So, a basic background in Linux at the level of usage of ECS 40 would be helpful but not required. Prior background in  and L^AT_EX themselves is NOT required.

3 Required Course Materials

The textbook is my open source book. You can buy it at the campus bookstore (printed at cost), or can download it at <http://heather.cs.ucdavis.edu/~matloff/132/PLN/ProbStatBookF2010.pdf>.

You are required to have a hard copy of the entire text.

4 Role of the Textbook

Make sure you bring the textbook to every lecture. It will play a very active role in the lectures.

Each day I will cover a portion of the textbook. I typically will ask you to open to a certain page, and then will call your attention to various sentences, Examples and pictures on that page. **The lectures will consist mainly of discussion and amplification of the Examples in the textbook, with additional Examples and comments.** I will seldom write on the blackboard.

SINCE THE TEXTBOOK ALREADY CONTAINS MY LECTURE, YOU DON'T TAKE NOTES (other than miscellaneous comments in the margin). THIS FREES YOUR ATTENTION TO ASK QUESTIONS AND TO ENGAGE IN CLASS DISCUSSION. You are very much encouraged to bring up any questions you might have. Please don't be shy about asking questions; there is no such thing as a "dumb" question. Similarly, listen carefully to the questions the other students ask; this can be a very valuable source of insight, to which you should pay special attention.

After each lecture, you should read in detail whatever portion of the notes was covered. Again, note that you should read all the details, as there is not enough time during lecture to discuss all of them.

5 Course Content

5.1 Topic Coverage

I tend to end up covering less material than I had hoped at the beginning of the quarter. So, the following is a goal, probably not fully achieved.

5.1.1 Review of Undergraduate Prerequisite Material

It will be presumed (with high probability, correctly) that you have forgotten what you learned in your probability prerequisite course. Thus in the beginning, there will be a semi-guided self-review, of the following chapters in the text:

Ch.1: all

Ch.2: all

Ch.3: all except Sec. 3.12

Ch.4: through Sec. 4.5

The self-review is "semi-guided" in the following sense. I will actually begin lecturing with Chapter 5. Most of that chapter is likely material not in your prerequisite course (or covered there just superficially). While

lecturing on Chapter 5, which is on multivariate distributions, I will constantly make comparisons to the univariate case, using that as review. Meanwhile, you will be reading the above chapters, doing homework in them that I assign, and taking one or more Quizzes on that material.

5.1.2 Core Course Topics

The core of the course will consist of:

advanced probability:

- Ch.5, multivariate probability models: through Sec. 5.11
- Ch.10, Markov chains: all
- Ch.11, renewal models: all
- Ch.12, queuing models: all

statistics/machine learning:

- Ch.6, simple confidence intervals and significance tests: all, but quickly, as most students have seen this, maybe even in high school
- Ch.7, advanced statistics: Secs. 7.1, 7.2, 7.3, 7.8
- Ch.8, model development and selection: all
- Ch.9, regression models and machine learning: all

5.2 What You Will Get Out of the Course

- Acquisition of valuable research tools: Many fields of CS and EE make heavy use of probability and statistics. After this course, you will have a good foundation from which to apply these tools, learn more about them and even develop your own new tools. You'll be able to read and understand statistical applications in the research literature, and even spot what's wrong with them (plenty, in many cases!).
- Impact on your daily life: Applications of probability and statistics show up everywhere—in election polls, in medical diagnosis, in insurance, in casinos, in the courtroom, and so on. What does all this mean, in a real, practical sense? Again, I'm not talking about formulas—the issue is, what does it really MEAN? You will get a pretty practical understanding of these concepts in our course.
- Acquisition of software tools: In this course, you will learn the basics of the R statistical programming language, which is the “gold standard” for data analysis, and \LaTeX , a word processing package that is standard in conferences and journals. For links to my tutorials on R and \LaTeX , see our course Web page, <http://heather.cs.ucdavis.edu/~matloff/256.html>.

6 Homework

6.1 Amount of Work

There will be about four or five written homework assignments during the quarter. These will consist mainly of mathematical analysis but sometimes will involve some light programming.

6.2 Writeups

Analytical work must be written in \LaTeX . See below.

All programs and graphs must be done in R.

6.3 Homework Partners

You are required do your homework together in groups, with group sizes to be announced (probably 3-4). Your group submits the homework together; make sure to include your group members' names in your writeup.

You must submit to me your group membership lists by 11:59 p.m. of the day of the third lecture. If you are not yet in a group, or have only a partial group, ask me to assign you to a group. Later, notify me immediately if, due to drops, your group size falls below the minimum group size.

In working with your partners, you must work TOGETHER, instead of simply dividing up the work among the partners.

6.4 Submitting Homework

In order to facilitate grading, I will require homework submission to be done exactly according to the following rules:

- Homework is due by 11:59 p.m. on the due date. Submit just one copy of your group's work, with all names on it. Submission is by e-mail to me.
- Title your e-mail message "HwknX," where n is the homework number in Roman numerals and X is a string consisting of the first four letters of each surname in your group (in alphabetical order). For instance, if the surnames of your group members are Crocker, Hopkins, Huntington and Stanford, then the title of your e-mail message for submitting Homework II would be HwkIICrocHopkHuntStan.
- Your message should consist of one attachment, a Unix **.tar** file, again with the name "HwknX.tar," in which the following are packaged:
 - Your writeup, in both **.tex** and **.pdf** files.
 - Any graphs, in **.pdf** files.
 - Any code, in **.r** files. Make sure to put your NAME in the comments!
 - A shell script (**bash** or **tcsh**) which runs your R code, if any.

Do NOT make a subdirectory. My automated shell script will apply **tar xf** to your **.tar** file and will then expect to see your files as listed above.

I apologize in advance for all this regimentation, but it will enable me to grade your homework more quickly and with better feedback. Thanks in advance for your cooperation.

7 Quizzes and Exam

In the past, I have just given an Exam on the last day of lecture. However, I found that this just caused students to fall behind. So in addition, I will also have three short Quizzes.

Each Quiz will be about 15 minutes long, and will consist of two or three short problems. I will still lecture on days on which I give Quizzes.

The Quizzes will be taken on an open-materials basis. Make sure to bring your textbook and any other written course materials, and you are welcome to bring in anything else you consider helpful, e.g. probability textbooks, English dictionaries, lucky charms, etc.

I will again give an Exam on the last day of class, occupying the entire class period. You will together with your homework group on the Exam.

Please make sure to bring your textbook to all Quizzes and the Exam. You will need it—many questions will refer to specific page numbers in the textbook.

8 Grading

The grading components will be as follows:

- Homework, 40%.
- Quizzes, 40%.
- Exam, 20%.

In recognition of the fact that you may be forced to miss a quiz due to illness and the like, I will drop your lowest quiz grade.

The threshold for a grade of n is $(n-1).85$; the threshold for an 'n-' grade is $(n-1).6$; the threshold for a '+' grade is $n.2$. So, for example, if your weighted average from the above formula is between 2.6 and 2.84, your course grade is B-; if the average is between 2.85 and 3.19, your course grade is B; if the average is between 3.2 and 3.59, your grade is B+.

9 Consultation

9.1 Office and Office Hours

My office is in 3053 Kemper, Ext. 2-1953. My office hours will be Mondays, 3:30-4:30 and Wednesdays, 4:30-5:30. **I enjoy my office hours very much, and look forward to interacting with you during them.** I am also available at other times if you have short questions.

9.2 E-Mail

You are welcome and encouraged to send me your questions via e-mail. I read my mail every day, including evenings and weekends.

10 The Class Blog

I will be posting announcements of various sorts (Homework, Exams, tips, etc.) on the class blog, <http://heather.cs.ucdavis.edu/xxx/Blog.html>, where xxx is our course number, e.g. <http://heather.cs.ucdavis.edu/132/Blog.html> for ECS 132.

The class blog will play a major role in our course. It is required that you read the blog every day. Not only will announcements of Homework assignments be posted there, but also there may be later corrections, hints, etc. And of course, dates of Exams and other related information will be announced on the blog too.

This is a no-comment blog, read-only instead of writable.. If you have a question for me, send it to me via e-mail. I will reply to you, and if I believe the topic is of general interest, I will post it to the blog.

11 Machines

11.1 Linux PCs in CSIF

Our base machines will be the Linux PCs in the CS instructional labs (CSIF). If you pre-enrolled in this course, or if you took a CS course last quarter, you should already have an account.

If you write code, it is not considered valid unless it works on CSIF.

Did you know that most computer graduates of UC campuses end up working in either a fully or partially Linux job? It is also worth mentioning that Intel has complained that “UCD grads don’t know Linux well enough.” (Macs are Unix-based, which is close enough, if you use the command line.)

In other words: If you’ve been using Windows to do your CS Homework, you’ve been limiting your future career opportunities! Not a smart strategy, right?

So, if you don’t know Linux, you should remedy that problem now, to enhance your future career prospects. And the only way to learn it is to USE it—not just for coursework, but also for e-mail, writing term papers, developing Web pages, video editing, and so on.

I have a Linux installation and usage tutorial at <http://heather.cs.ucdavis.edu/~matloff/Linux/LinuxInstall.pdf>. If you are going to use Linux in your daily life, as I urge you to do, then I recommend installing on a dual-boot basis. Otherwise, install on USB key or an external hard drive. Instructions are given in my tutorial.

12 Analytical Work Must Be Done in L^AT_EX

Most of the Homework questions will be mathematical in nature. It is REQUIRED that you write these up in the L^AT_EX word processing package. L^AT_EX is widely used in engineering and the physical sciences.

Did you know that L^AT_EX is used as the math typesetting vehicle in



You can learn enough L^AT_EX for this course in about 10 minutes; start at <http://heather.cs.ucdavis.edu/~matloff/latex.html>.

Submit BOTH your **.tex** source file AND your **.pdf** file with your Homework.

It is expected that you do a professional job on your writeup. It should look clean and organized. You should explain your steps in prose; **don't just write a bunch of equations without explanation**. Use professional notation. For instance, the use of * for multiplication should be limited to program code, and should not be used in equations, as it is not math notation. Use juxtaposition or the $\text{\LaTeX}\backslash\text{cdot}$ instead.