

STATISTICS

Welcome to Math 3200! My name is Professor Edward Spitznagel. This is the successor course to Math 320. It is a calculus-based introductory course in statistics and the underlying probability theory supporting it. Since this course is now differentiated (and integrated—) from the effectively non-calculus-based Math 2200, a paragraph or two of explanation is warranted.

When I began teaching Math 320 in 1970, it had an enrollment of 21 students. At that time, it was a calculus-based course. Over the years, it grew until several years ago it had over 400 students. Gradually, the calculus prerequisite became a nominal one-semester dose (Math 131), which meant that the quality of the course really suffered. Perhaps that would not have been a problem, except for the fact that many of our upper level courses depended on students being prepared for them by Math 320. Without that preparation, Those courses had to spend their first third in reviewing what should have been covered in Math 320, and thus themselves became watered down.

By returning Math 320 to its roots, we hope to upgrade the quality of all our statistics offerings for both mathematics majors and minors. Of course, any student, major, minor, or not, who has the calculus background is welcome in the revitalized Math 320. Although what we are doing is in fact restoring Math 320 to what it once was, it was decided that it might be more politically correct to give it a new number—thus its new designation as Math 3200.

Times and Places

Our course meets Monday, Wednesday, and Friday 9-10 in Busch 100. **Before you come to class, please preview the section of the book to be covered that day.** Naturally I don't expect you to learn all the material from that reading. What I do expect is that you will be able to ask much better questions, having done that preview.

My *official* office hours are from 8:00 to 8:45 on Monday, Wednesday, and Friday in Busch 100. After my three classes end at noon, I typically take the Metro down to the Med School. You might find me in Room 118 of Cupples I when I get back late in the afternoon. You are *welcome* to knock anytime you see the light on. However, I do recommend calling in advance to see if I'm in. My telephone number is 935-6745.

Textbook

The text is Tamhane and Dunlop *Statistics and Data Analysis: From Elementary to Intermediate*. This is one of a very few books from which a junior level course can be taught. Most other books are either too hard (too much mathematics) or too soft (too little mathematics). Like Baby Bear's bed, Tamhane and Dunlop is just right. I have to confess that I did use the book once before, for all of the previous Math 320, and there was a lot of kvetching about it. It seems to have been a matter of *μαργαριτας εμπροσθεν χοιρον*, which I don't think will apply here.

Hand Held Technology

The Texas Instruments calculators TI-83, TI-84, and TI-89 (and the new TI-Nspire series) contain essentially every probability function and statistical program we will be using during the course. I have declared the above to be the official calculators for the course. I have a computer emulation of the TI-83, with which I will frequently work problems in class, projecting an image of the calculator on the screen. These calculators contain probability functions that are more accurate and easier to use than the distribution tables in the back of the book. Therefore I will not provide those tables for the examinations; you will be expected to use the calculator instead. *Verbum sapienti!*

Manual Homework

There are (usually) six recommended homework problems per day of class. In most instances, two are odd-numbered, with answers in the back of the book. The other four are even-numbered book problems. I will usually have time to work two even-numbered problems in class, leaving you with a net four problems per day to do on your own. These problems will not be graded. Your primary motivation for keeping up with the homework is that most of the examination problems will be homework problems with simple changes in the data.

For those of you who wish it, a grader will provide you with feedback via email on any problems you choose to do. Those who participated regularly in this service last year all achieved course grades of A- or higher. By 9AM of the Tuesdays and Thursdays following the Monday and Wednesday classes, you may drop off your solutions of whatever problems you wish in the Math Dept office, Room 100 of Cupples I. Following the Friday class, you may drop your solutions in the Room 100 door's mail slot by 9AM Monday. (Please don't give it to me in class Monday because there's too much risk it will get mixed up with

the required homework being handed in then.)

Please write only on the front side of each page, use a paperclip (not a staple) to hold them together, and pull off any jaggies if you tore them out of a notebook. Print the course number (Math 3200) and your email address *clearly* at the top of each page. We will score your solutions and email you scanned copies.

For those of you studying as a team, just submit one copy. Whoever submits it will receive the email and can forward it to everyone else. We're sorry that, due to the limitations of our scanner, we can only email a scored assignment back to a single address.

There are three simple conditions on this offer. First, we will only score original, hand-written work, not photocopies. Second, we will only score good-faith attempts to solve the problems; we will not write in solutions, or even provide answers, on blank sheets of paper. Third, we will not score illegible solutions; we will simply return these marked as illegible.

We will keep no records of how well you did on these problems. This is strictly a feedback service. There is no need to give us your name; just provide your email address.

Computer Technology

There is a wide variety of computer software for doing statistics, ranging from the relatively primitive capabilities in Microsoft Excel® to the extremely powerful SAS® package. We will use four statistics packages: SAS, STATA®, R®, and SPSS®. I introduce you to these four main packages because all are in common use, and you can never tell which one(s) might get you that envied job or internship. Recently a student wrote me the following thank-you note: "I'm starting an internship with the data analytic department at _____ next semester, and they were extremely impressed that I already had an introduction to SAS and SPSS. I don't

think I would even have had a chance at that internship if I didn't take the class."

I will demonstrate all four in class, and will assign homework problems for you to do and hand in for grading. The primary package will be SAS. We will cover the others in compare-and-contrast mode, so that you will be able to claim at least passing familiarity with all four when the time comes to interview for jobs and internships. The ArtSci laboratory in the basement of Seigle Hall has SAS, STATA, and R installed on its computers. When the time comes, I will introduce you to SPSS via a freeware clone called PSPP.

Computer Homework

There is one required computer homework exercise per week of class. In past years these exercises were chosen from the recommended manual homework problems. This year we will do something different, in which each week you will do "case studies" using large public data sets, such as the 2010 Census, the NORC General Social Survey or NHANES. I think you will have more fun analyzing real data, and you may be able to use that data in other courses. Each week I will send you an e-mail message explaining how to access the data. These exercises are due in class each Monday, with the exception of Labor Day and the Mondays immediately following Fall Break and Thanksgiving. That works out to a total of ten assignments. The computer homework will count as 20% of your course grade.

Unfortunately, the number of computers publicly available to Arts and Science undergraduates has dwindled from more than 60 down to the 14 currently in Seigle Hall Room L012: 10 PC's and 4 Mac's. In earlier years, I used to drop by Sunday afternoons to the large and spacious lab beneath Holmes Lounge. Typically more than half the class would show up at that time to do their homework. My eagle eye would help them spot missing semicolons, unbalanced quotes, and other little details that can drive a person to drink. Regrettably,

with the small size of the Seigle Lab, that is no longer feasible.

Thankfully, though, Seigle L012 is open very generous hours:

Monday: 8:30am-8:30pm
Tuesday: 8:30am-8:30pm
Wednesday: 8:30am-8:30pm
Thursday: 8:30am-8:30pm
Friday: 8:30am-5:30pm
Saturday: 11:00am-5:00pm
Sunday: 11:00am-5:00pm

As long as you don't all wait until late Sunday afternoon, you should be ok. As far as dealing with those mischievous semicolons and quotation marks, feel free to bring snippets of code (printed on paper or copied to flash drive) to my office hours. I will have the appropriate software running on the computer for you to show me what is bugging you. Also, if you get the feeling that 14 public computers are insufficient for more than 4000 A&S undergraduates, you might mention to your academic advisors that you expect a little more for the tuition you are paying.

Examinations

As mentioned earlier, examinations are closely linked to the homework problems. If you faithfully work the problems, you should have no trouble scoring well on the examinations. Each examination will contain twenty multiple choice problems, of which approximately fifteen will be homework problems with altered numbers. You may bring one 4×6 inch notecard to each in-semester examination. For the final exam, you will be permitted to bring all your previous notecards, plus one you have prepared for the final exam. You may use both sides of each note-card.

Over the four examinations, you can achieve a maximum of 80 points. With the computer homework added in, your maximum number of points will be 100. At the end of the semester, the A range will be 90 and above, the B range will be 80 to 90, the C range will be 70 to 80, and the D range will be 60 to 70, with plus and minus grades at the tops and bottoms of each of these ranges.

Students ask if I ever grade on a “curve.” Curve grading was popular about fifty years ago. It assigned six letter grades A, B, C, D, E, and F based on a Gaussian, also called a “normal” curve. The grade of A corresponded to being 2 standard deviations above the mean and was awarded to the upper 2.5% of all students. The grade of B corresponded to being one to two standard deviations above the mean and was awarded to 13.6% of all students. The most common grades were C and D, at 34.1% each. I doubt any of you would like the grades to be assigned based on that system.

Instead, I will follow the modern convention, in which the A range will be 90 to 100, the B range will be 80 to 90, the C range will be 70 to 80, and the D range will be 60 to 70, with plus and minus grades at the tops and bottoms of each of these ranges. If you are registered pass/fail, you must achieve at least 70 points to pass, which is the lowest score for a C-.)

In addition to calculating the straight sum of points, I will also average the examination scores using a weighting process, in which each in-semester examination counts 16% and the final counts 32%, giving you whichever score is higher. (The computer homework will still be counted at 20%.)

This alternative weighting system rewards students who have tended to improve over the semester.

Examination Schedule

The three in-semester examinations will be given from 7PM to 9PM the following **Wednesday evenings**: September 18th, October 16th, and November 13th.

The final examination will be given on **Thursday, December 12, 3:30PM-5:30PM**.

As always, examination room assignments are posted on the Math Dept website:

<http://www.math.wustl.edu/seatlookup/>

the day of the examination.

Recommended Homework

Following are the recommended homework problems. At the risk of preaching to the choir, let me say that mastering these and reading the book should give you the traditional two-hours-out-of-class-for-every-one-in-class needed for success in the typical undergraduate course. The last time I taught a course (differential equations) to engineering students, they complained to their dean that I was working them too hard, giving them homework that took two hours (gasp, shudder) per class period. He asked them how much homework time they spent in their other courses, and they said, oh, about half an hour. Would you really want to fly in an airplane designed by engineers like that?

Two top-notch universities, CalTech and MIT, award credits equal to the weekly sum of lecture hours and expected amount of hours outside of class. As a reality check, I visited their websites and found the credits for their equivalent statistics courses to be:

CalTech: Ma112a lists 9 units of credit.
MIT: 18.443 lists 12 units of credit.

Thus, these two schools expect their students to spend between two and three hours outside of class for every hour inside class.

Aug 28	Chapter 2	4,5,6,10,12,15
Aug 30	Chapter 2	16,17,18,20,26,27
Sept 2	Labor Day Holiday	
Sept 4	Chapter 2	28,29,30,32,33,34
Sept 6	Chapter 2	35,36,38,40,41,42
Sept 9	Chapter 2	46,48,49,52,53,54
Sept 11	Chapter 2	61,62,63,64,66,70
Sept 13	Chapter 2	71,72,74,75,76,77
Sept 16	Chapter 2	78,79,80,81,82,83
Sept 18	Chapter 3	1,2,3,4,5,6
Sept 18	First Examination	
Sept 20	Chapter 3	7,8,9,10,11
Sept 23	Chapter 3	12,14,15,16,17,18
Sept 25	Chapter 3	19,20,21,22,24,26
Sept 27	Chapter 4	1,2,4,5,7,8
Sept 30	Chapter 4	10,11,12,13,18,22
Oct 2	Chapter 4	30,32,33,34,36,39
Oct 4	Chapter 5	2,3,4,6,11,12
Oct 7	Chapter 5	16,17,18,20,22,23

Oct 9	Chapter 5	24,25,26,29,30,33	Nov 13	Chapter 11	22,23,28,30,34,37
Oct 11	Chapter 6	1,2,4,6,7,8	Nov 13	Third Examination	
Oct 14	Chapter 6	11,12,13,14,15,16	Nov 15	Chapter 12.	1,2,3,4,6,7
Oct 16	Chapter 6	17,18,21,22,26,28	Nov 18	Chapter 12	8,10,11,12,13,16
Oct 16	Second Examination		Nov 20	Chapter 12	18,19,20,21,22,24
Oct 18	Fall Break		Nov 22	Chapter 13	1,3,6,16,20,22
Oct 21	Chapter 7	1,6,8,12,14,15	Nov 25	Chapter 13	25,26,28,29,30,35
Oct 23	Chapter 7	17,18,19,20,22,25	Nov 27	Thanksgiving Holiday	
Oct 25	Chapter 8	1,2,3,5,6,8	Nov 29	Thanksgiving Holiday	
Oct 28	Chapter 8	9,10,11,16,18,22	Dec 2	Chapter 14	2,3,4,11,14,16
Oct 30	Chapter 9	2,6,7,12,14,15	Dec 4	Chapter 14	19,20,21,23,24,26
Nov 1	Chapter 9	17,20,24,28,34,41	Dec 6	Chapter 15	14,15,16,17
Nov 4	Chapter 10	2,4,5,6,7,8	Dec 9-11	Reading Period	
Nov 6	Chapter 10	9,10,11,16,20,24	Dec 12	Final Examination	
Nov 8	Chapter 10	28,29,30,31,32,36			
Nov 11	Chapter 11	2,3,6,10,12,17			