

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 three 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 12 to 1 on Monday and Wednesday, in Room 118 of Cupples I. However, I am there at other times, on average perhaps 80 hours a week (Me, a workaholic?), and you are *welcome* to knock anytime you see the light on. 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 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 contain essentially every probability function and statistical program we will be using during the course. It would be foolish not to use such technology in our course, as it saves memorizing a huge number of arcane formulas. I have therefore 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 also contain functions that supersede the distribution tables in the back of the book. I will not provide those tables for the examinations; you will be expected to use the calculator instead. *Verbum sapienti.*

Manual Homework

There are six recommended homework problems per day of class. Two are odd-numbered, with answers in the back of the book. The other four are even-numbered. I will usually work two of these even-numbered problems in classes, 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 slip your solutions under my door, Room 118 of Cupples I, by noon Saturday.

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 your Washington University 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 two simple conditions on this offer. First, we will only score original handwritten work, not photocopies. Second, we will only score good-faith attempts to solve the problems. We will not write solutions, or even provide answers, on blank sheets of paper.

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 will demonstrate all of them 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 Eads Hall has SAS, STATA, and R installed on all 38 of its PC's. I have bundled the one that isn't, SPSS, with your textbook.

Computer Homework

There are three required computer homework problems per week of class. When it is convenient, these problems are chosen from the recommended manual homework problems. These problems are due in class each Monday, with the exception of Labor Day (no class), the Monday following Fall Break, the Monday following Thanksgiving, and the last day of class. That works out to a total of eleven assignments. Each Sunday before an assignment is due, I will drop down to the ArtSci computer lab from 4pm to 6pm, to see if I can be of assistance. The computer homework will count as 20% of your course grade.

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 partial-credit graded problems, of which approximately fifteen will be homework problems with altered numbers. You may bring one 4×6 inch notecard to each examination.

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 following 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%.)

The latter weighting system rewards students who have tended to improve over the semester.

Examination Schedule

The three in-semester examinations will be given from 6:30PM to 8:30PM on the following Tuesday evenings: September 22nd, October 20th, and November 17th.

The final examination will be given on **Friday, December 11th, 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. In each day’s list, two or occasionally three are odd-numbered. You will find answers for those in the back of the book.

At the risk of preaching to the choir, let me say that mastering these and reading the book should give you the two hours-out-of-class-for-every-one-in-class needed for success in the typical undergraduate course.

Two schools, 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 26	Chapter 2	6,9,10,11,12,14
Aug 28	Chapter 2	16,17,18,20,22,25
Aug 31	Chapter 2	28,29,30,32,33,34
Sept 2	Chapter 2	36,39,40,41,42,46
Sept 4	Chapter 2	48,49,50,52,53,54
Sept 7	Labor Day Holiday	
Sept 9	Chapter 2	59,60,61,62,64,70
Sept 11	Chapter 2	71,72,73,74,75,76
Sept 14	Chapter 2	78,79,80,81,82,83
Sept 16	Chapter 3	1,2,3,4,5,6
Sept 18	Chapter 3	7,8,9,10,11
Sept 21	Chapter 3	12,14,15,16,17,18
Sept 22	First Examination	
Sept 23	Chapter 3	20,21,22,23,24,26
Sept 25	Chapter 4	2,3,4,5,7,8
Sept 28	Chapter 4	9,10,11,12,14,28
Sept 30	Chapter 4	30,31,33,34,38,40
Oct 2	Chapter 5	2,3,4,6,7,8
Oct 5	Chapter 5	16,18,19,20,22,23
Oct 7	Chapter 5	24,25,26,27,30,32
Oct 9	Chapter 6	1,2,3,4,7,8
Oct 12	Chapter 6	11,12,13,14,15,16
Oct 14	Chapter 6	17,18,20,22,24,30
Oct 16	Fall Break	
Oct 19	Chapter 7	1,7,8,12,13,16
Oct 20	Second Examination	
Oct 21	Chapter 7	17,18,19,20,21,22
Oct 23	Chapter 8	1,2,3,6,7,8
Oct 26	Chapter 8	9,10,13,16,18,20
Oct 28	Chapter 9	5,6,8,11,14,16

Oct 30	Chapter 9	17,20,22,27,28,32
Nov 2	Chapter 10	2,4,5,6,7,8
Nov 4	Chapter 10	9,10,15,16,20,24
Nov 6	Chapter 10	28,29,30,31,32,34
Nov 9	Chapter 11	2,3,4,11,12,17
Nov 11	Chapter 11	22,23,28,30,34,37
Nov 13	Chapter 11	40,41,42,44,45,46
Nov 16	Chapter 12.	1,2,3,4,5,6
Nov 17	Third Examination	
Nov 18	Chapter 12	8,9,10,11,12,13
Nov 20	Chapter 12	18,19,20,21,24,28
Nov 23	Chapter 13	2,3,6,16,17,22
Nov 25	Thanksgiving Holiday	
Nov 27	Thanksgiving Holiday	
Nov 30	Chapter 13	25,26,28,29,30,34
Dec 2	Chapter 14	2,3,4,12,13,16
Dec 4	Chapter 14	19,20,21,23,24,25
Dec 7	Chapter 15	14,15,16,17
Dec 8-10	Reading Period	
Dec 11	Final Examination	

Required Homework

Here are the required computer homework problems. Three problems are due per week, always on Monday, at the beginning of class. Three Mondays are skipped, making the total number of assignments equal to twelve. All assignments are to be done with SAS. In addition, most assignments are also to be done with STATA, R, or SPSS, on a semi-rotating basis. I will let you know on a week-by-week basis what other package is to be used that week, in addition to SAS.

Aug 31	2.20(simulate), 2.22, 2.25
Sept 14	2.39(simulate), 2.60, 2.73
Sept 21	2.81, 2.82, 3.10(simulate)
Sept 28	3.14($n=300$), 4.7, 4.8
Oct 5	4.14, 4.34, 5.3
Oct 12	5.20, 5.27, 6.2ab(simulate)
Oct 26	6.12, 7.16, 8.7
Nov 2	8.16, 9.6, 9.28
Nov 9	10.5, 10.20, 10.30
Nov 16	11.2, 11.22, 11.37
Nov 23	12.2, 12.13, 12.24
Dec 7	13.6, 13.30, 14.12