Statistics 212 – Principles of Statistics II

Class Information
Section 502, Spring 2014

Statistics 212 is an introduction to standard statistical methods such as multiple linear regression, experimental design and the analysis of variance, and categorical data analysis. Advanced methods that also may be discussed include general linear models, nonparametric density estimation, nonlinear and nonparametric regression, logistic regression, and inference for percentiles.

Time and Place: MWF 1:50pm–2:40pm, Blocker 169.
Instructor: Prof. Daren Cline, Blocker 459D, 845-1443. e-mail: dcline@stat.tamu.edu
Office Hours: MWF 10:00am–11:00am or by appointment.
Prerequisite: Statistics 211 or equivalent (calculus-based introduction to statistics).
Grader: Sha Ni, e-mail: sn907@stat.tamu.edu
Office hours: Tues 10:00–11:00 am, Thurs 12:00–1:00, Blocker 405A or 405F. You may contact the grader for help on homework or questions on graded homework. The grader will not accept late homework under any circumstances. Requests for late homework (with valid university excuses) and questions about the exams should be made directly to me.

Help Sessions: Wed, Thu, Sun 1:00pm–3:00pm, Blocker 162. TAs (not just your grader) will be available to help you. The laptops in the room are for statistics class use only.

Online Resource: eCampus (See below for more details.)

Course Objectives: The purpose of statistics is to provide defensible analyses of experimental data. In this class you will be expected to
• Understand models relating a response variable to one or more other variables.
• Know the basic forms of experimental design and their purposes.
• Know the correct statistical analyses, their purposes and limitations, what to conclude and what not to conclude.
• Use a statistical application to analyze data.
• Understand what makes a conclusion defensible.
• Understand assumptions and how to check them.

Notes and Handouts: You are expected to download lecture notes and other handouts as they become available on eCampus, and to bring them to class. The handouts include output from statistical computing packages with some explanation about how they were obtained.
Discussion Board: A discussion board (in eCampus), monitored by the grader and myself, is available at the DoStat website. I will originate topics as I see fit or by request; you may start threads within topics. Its purpose is to allow you to converse freely with others in the class about the course, especially for homework. I only ask that you give each other help in the form of hints and suggestions, but not complete solutions. Courtesy and discretion are of course required.

*The discussion board will be off-limits on exam days.*

Computing: I will provide examples using JMP and other statistical software, which you may refer to for doing homework (recommended). JMP will be made available to you for free (for *your own computer*) after the start of the semester. The computers in the Open Access Labs have JMP also.

Disabilities Help: The Americans with Disabilities Act ensures that students with disabilities have reasonable accommodation in their learning environment. If you have a disability and need help, please contact me and Disability Services in B118 Cain Hall, 845-1637.

Academic Integrity: You are expected to follow the Aggie Honor Code and maintain the highest integrity in your work for this class. This includes not passing off anyone else's work as your own, even with their permission. Please see the homework and exam policies below for specifics.

Copyright: All the resources I provide for this course are copyrighted and may not be copied, sold or distributed without my express, written permission.

Homework: Homework is worth 20% of the total term score. None may be dropped. It will be assigned in eCampus and collected regularly.

*Late homework will not be accepted without an approved excuse.*

Method and communication are as important in this course as are final solutions. Homework is to be detailed and clear, with all steps provided, on 8½×11 paper and stapled in the upper left corner. Computer output should be pasted into solutions as needed. (Do not append it to the end of your homework.)

*Please see the homework policy below.*

Exams: There will be two midterm quizzes worth 22.5% each and a final worth 35%. All exams are cumulative and closed book. Old exams will be made available for review.

Bring a large (blue/gray) scantron sheet. You will be allowed to bring statistical tables and one additional page (8½×11, both sides) of notes and formulas per exam.

*Please see the exam policy below.*

Exam Dates: Exam I: Friday, 20 February. *(subject to change)*
Exam II: Wednesday, 1 April. *(subject to change)*
Final Exam: Monday, 11 May, 3:30pm–5:30pm.
Statistics 212 – Principles of Statistics II

Grading Scale:

A: 85% – 100%
B: 75% – 84%
C: 60% – 74%
D: 50% – 59%

Important Dates:

Wednesday, 21 January – first day of class.
Monday, 26 January – last day to add/drop.
Monday, 16 March – first day of spring break.
Friday, 3 April – no class (reading day).
Tuesday, 21 April – last day to Q-drop or withdraw.
Tuesday, 5 May – last day of class (redefined Friday).

eCampus Instructions:

All students will have access to the lecture notes and handouts, homework assignments and other instructional materials at the eCampus web site.
1. Go to ecampus.tamu.edu.
2. Click on Log In and log in with your NetID.
4. Click on Discussions in the menu to the left to access the discussion board.
5. Click on Content in the menu to the left for lecture notes, handouts, homework assignments and any other resources I may provide.

Homework Policy:

Homework assignments will be downloaded from eCampus. Your homework solutions must be your own work, not from outside sources, consistent with the university rules on academic integrity. I expect you to follow this policy scrupulously. Your exam performance is much more likely to be better.

You may use:

• Your textbook, e-book, and notes from this class.
• Your notes, homework, etc., from a related class that you took or are taking.
• Discussion with the instructor or grader.
• Voluntary, mutual and cooperative discussion with other students currently taking the class. Do not post solutions (anywhere). Suggestions, descriptions and partial explanations are ok.

You may not use:

• Solutions manuals (printed or electronic) other than the student manual.
• Solutions from students who took the class previously.
• Simply copying from students in this class, including expecting them to reveal or provide their solutions in "discussion". That is, you may work together as indicated above as long as you prepare your own solutions.

Homework is to be submitted by its due date unless I specify otherwise. Late homework is not acceptable.
Exam Policy:

Exams will be comprehensive, cumulative and closed book. *Your exam solutions and answers must be your own work*, consistent with the university rules on academic integrity. Bring a large (blue/gray) scantron sheet for your answers. Acceptable resources are:

- A calculator for numerical calculations only. The calculator *may not be part of, associated with or connected to any communication device*, such as a cell phone, iPod, tablet or laptop.
- Statistical tables. ([Bring your own copies. I will have versions available online.](#))
- One page (8½×11, both sides) of notes for the first exam, two pages for the second exam and four pages for the final exam. These *must be of your own construction*, not copied from somewhere or someone else.

No other resources are acceptable. *The discussion board will be off-limits on exam days.*

*No exam may be taken early or made up*, except if you provide a university excused absence with appropriate documentation.

Copies of old exams will available to you for review. However, their content may not exactly match this semester's exams.

Classroom:

*Please turn off all communication devices* (cell phones, iPods, etc.) while in the classroom. You can have a calculator for in-class work. A laptop or tablet is ok as long as you only use it to take notes or to view lecture notes and handouts for this course.

Questions are encouraged, especially to help clarify points in the lecture. No question is "bad" or "dumb" if it is relevant (although I do appreciate it if you listen and avoid asking a question just answered).

Help Session Lab Room:

You are encouraged to take advantage of the help sessions in Blocker 162. Various TAs will be available, depending on the time, but all are familiar with this and similar courses. *The laptops in the room are for statistics class use only.* Please be considerate of others.

Missed Work and Incompletes:

- If you must miss an exam due to illness or circumstances beyond your control, notify me or the Statistics Department, in writing or by email (before, if feasible, otherwise within two working days after you return). See me as soon as possible to schedule a make-up exam.
- An incomplete will be given only in the event that circumstances beyond your control cause prolonged absence from class and the work cannot be made up.
## Course Outline

**Topic** | **Chapters**
--- | ---
**I. Estimating Distributions** | 1, 4, 8
  a. histogram, density plot, box-plot (review) | 1.1–1.4, 4.1
  b. cumulative distribution, quantile plot (review) | 4.2, 4.6
  c. hypothesis testing (review) | 8.1, 8.2, 8.4, 8.5
**II. Regression and Correlation** | 5, 12, 13
  a. correlation and conditional expectation (review) | 5.1–5.2
  b. straight line regression (review) | 12.1–12.4
  c. inference for correlation | 12.5
  d. checking for violations of assumptions | 13.1
  e. polynomial and nonparametric regression, transformations | 13.2–13.3
  f. multiple linear regression | 13.4
  g. model selection and other issues | 13.5
**III. Design and Analysis of Experiments** | 9, 10, 11, 15
  a. completely randomized design (review) | 10.1
  b. multiple comparisons and contrasts (review) | 9.2, 10.2
  c. assumptions, transformations and Kruskal-Wallis test | 9.3, 10.3, 15.2–15.4
  d. randomized block design, Friedman test | 11.1, 15.4
  e. factorial experiments and interaction | 11.2–11.4
  f. random and mixed effect models | 10.3, 11.2
  g. general linear models, covariate analysis | 
**IV. Analysis of Categorical and Count Data** | 2, 3, 5, 8, 9, 13, 14
  a. one and two sample binomial procedures (review) | 3.4, 3.6, 8.3, 9.4
  b. multinomial experiments | 5.1, 14.1
  c. chi-squared goodness of fit test | 14.1–14.2
  d. conditional probability, independence (review) | 2.4–2.5
  e. contingency test, homogeneity test, McNemar test | 14.3
  f. logistic regression | 13.2
**V. Methods for Percentiles** | 15
  a. sign test, tests for percentiles | 15.1
  b. confidence intervals | 
  c. median regression | 