

Stat 1040: Introduction to Statistics

Syllabus

Fall Semester, 2014

Tuesday and Thursday, 9:00-10:45 in ESLC 130

Introduction

Statistics is the collection, display, and analysis of data; it is the art of making wise decisions in the face of uncertainty. The purpose of this class is to teach you how to think critically about data—how it was collected and analyzed—and its uses in addressing interesting questions. Some of the issues we will touch on are:

- How should experiments be designed to measure the effects of new treatments (medical or other)?
- What causes the resemblance between parents and children, and how strong is that force?
- Why does a casino make a profit at roulette?
- How can the Gallup poll predict election results so accurately using samples of only a few thousand people?
- How can we tell if medical treatments or other interventions really work or if the results we observe are simply due to chance error?
- How can we tell if sex bias or race bias has occurred?

You will also learn about some interesting applications of probability models to, for example, genetics, and how the great statistician R.A. Fisher was able to show that the results of some of the earliest experiments in genetics were probably “made-up.”

Instructor

Robert Heal, Professor

Department of Mathematics and Statistics

Office: 306 AnSci, Phone: 760-5771

Office Hours: Tuesday, Thursday 2:30pm—3:30pm

Email: Robert.Heal@usu.edu Webpage: <http://www.math.usu.edu/~rheal/webpage>

Textbook

Statistics, by Freedman, Pisani and Purves, Fourth Edition (W.W. Norton, 2007). We will cover the entire book except Chapter 11 sections 4 and 5, Chapter 15, Chapter 22, Chapter 24, Chapter 28 section 3, and the finite population correction factor.

Calculator

A calculator is required; the **CASIO** *fx-260SOLAR* is recommended for this class and is available at the bookstore.

Recitations

All students must be registered for a recitation section. Recitations meet on Wednesdays and Fridays. Students are expected to attend lectures and recitations. Recitations will be devoted to working problems and answering questions.

Classroom Behavior

Please turn off your cell phones as a courtesy to the instructor and your classmates. No texting, web browsing or electronic games in class. No communication devices, music players, etc may be used in tests. No baseball caps in tests. If you use a laptop to view class material during the lecture, please sit at the back or on one side of the classroom so that you minimize any distraction for other students.

Grades

Your grade will be based on a score out of 500, made up of 100 points for quizzes, 100 points for each midterm, and 200 points for the final exam. An approximate grade distribution is given below. Plus and minus grades are not indicated.

20% A

25% B

45% C

10% D, F

Tutoring

Free tutoring for Stat 1040 is provided by the Academic Resources Center. The schedule is available at <http://www.usu.edu/arc/tutoring/>

Quizzes

There are eleven weekly quizzes (10-points each) that you will take at the beginning of a recitation meeting. Your lowest quiz score will be dropped. You will **not** be allowed to take the quizzes at a different time. Only USU sponsored events, jury duty, military duty, serious illness or death in the family are acceptable excuse and must be supported in writing.

Midterms

There will be two midterms. Both midterms will be closed book. You need to bring a calculator to the midterms. Only USU sponsored events, jury duty, military duty,

serious illness or death in the family are acceptable excuse and must be supported in writing.

First Midterm: Chapters 1-12 (inclusive) September 30, in class.

Second Midterm: Chapters 13-23 (inclusive) November 6, in class.

Final Exam

The final exam is closed book and comprehensive, but it will emphasize the material in chapters 26 through 28. The final exam is scheduled to be in ESLC 130 on:

Thursday December 11 from 9:30am to 11:20pm.

If you think you may be unable to attend the final exam for any reason, you should drop this class now and take it in a semester when you will be able to attend the final. You will NOT be allowed to take the final at a different time because of family weddings, birthdays, anniversaries or because of vacation plans. In the case of a last-minute medical emergency, or a death in the family, you will be given an incomplete and required to take the common final exam the next time the class is offered.

Please note: Qualified students with disabilities may be eligible for reasonable accommodations. All accommodations are coordinated through the Disability Resource Center (DRC) in Room 101 of the University Inn, 797-2444 voice, 797-0740 TTY, or toll free at 1-800-259-2966. Please contact the DRC as early in the semester as possible. Alternate format materials (Braille, large print or digital) are available with advance notice.

Reading

Each week you have a reading assignment consisting of chapters from the book. This material will be discussed in the lectures. Part of your reading responsibility is doing the exercise sets at the end of each section. Your ability to do these is a measure of your understanding of what you have read. Consequently, a suggested reading program for each chapter is as follows:

1. Read the **summary** at the end of the chapter, taking careful notice of the key words that are in *italic* type.
2. Read the **entire chapter** without looking at the exercise sets, but paying attention to the key words.
3. Reread the chapter, this time doing the **exercise sets**. Whenever necessary, look back at the chapter to find the ideas that will help you to do the exercises. Check your answers against those in the back of the book when you have finished a problem. Do not get in the habit of looking at the answers before you attempt the problems!
4. Reread the summary to make sure you now understand the **key words**.

Topics

1. *Design of experiments*: controls, randomization, blind and double blind, placebos. The Salk vaccine trial, historical controls.
2. *Observational studies*: association and causation, confounding factors, the clofibrate trial, sex-bias in graduate admissions.
3. *Descriptive statistics*: the histogram, the density scale, cross-tabulation, the average and the SD and their relationship to the histogram, the median, the normal approximation for data, percentiles, percentiles and the normal curve, measurement error, outliers, bias versus chance error, plotting points and lines, the slope and the intercept of a line.
4. *Correlation*: the scatter diagram, the correlation coefficient, properties of the correlation coefficient (invariance to change of location and scale, symmetry), ecological correlations, correlation does not imply causation, examples, the SD line.
5. *Regression*: the graph of averages, regression to the mean, the regression method for individuals, the regression fallacy, there are two regression lines, the r.m.s. error, plotting residuals, heteroscedasticity, looking at vertical strips, using the normal curve inside vertical strips, the slope and the intercept, the method of least squares, does regression make sense?
6. *Probability*: the long run argument, conditional probabilities, the multiplication rule, independence, the Collins case, listing the ways, the addition rule the Paradox of Chevalier de Mere, are real dice fair? The binomial formula.
7. *Chance variability*: What does the law of averages really say? Chance processes, the sum of the draws, making box models, the expected value and the standard error, using the normal curve, probability histograms, the normal approximation for probability histograms and its scope.
8. *Sampling*: The Literary Digest poll, the year the polls elected Dewey, using chance in survey work, how well do probability methods work, a close look at the Gallup poll, telephone surveys, chance error, bias, quota samples, samples of convenience.
9. *Chance errors in sampling*: the standard error, the correction factor, the Gallup poll, the accuracy of percentages, confidence intervals and their interpretation, the accuracy of averages, the SE and expected value of an average.

10. *Tests of significance*: null and alternative hypotheses, test statistics and significance levels, the role of the box model, zero-one boxes, the one-sample z-tests and *t*-tests, the standard error for a difference, comparing two sample averages, comparing two proportions, experiments, the χ^2 -tests for independence and goodness-of-fit, how Fisher used the χ^2 -test to show that Mendel's data was fudged, was the result significant? Was it important? Data snooping, the importance of the box model.