Math 207, Probability and Statistics
Spring Semester, 2007
Dr. Evelyn Bailey

Office hours: Room 122 Pierce Hall, posted weekly on the class conference

Text: Introduction to Mathematical Statistics, 5th ed, Paul G. Hoel

Reader: Statistics You Can't Trust, Steve Campbell

Materials: Math 207 Notes (provided in a notebook for this class), use of a calculator (TI-83, TI-83 PLUS, TI-84, or equivalent type)

Content: Visual displays of data, measures of central tendency and of variability, classification of data, counting, probability, Bayes Theorem, probability functions, Chebyshev's Theorem, discrete distributions (binomial, hypergeometric, Poisson, uniform), continuous distributions (Exponential, Normal, Uniform), Central Limit Theorem, Confidence Intervals, Hypothesis Testing (for means and for proportions, for one and two samples), goodness of fit, Chi Square Contingency Tables, one-way ANOVA, simple linear regression and correlation, nonparametric tests (median, Wilcoxon Rank Sum Test, Kruskal-Wallis Test).

Goals: At the end of this course each student should be able to: categorize data, work various probability problems, understand the role of functions in statistics, describe major misuses of statistics, recognize several standard distributions, analyze interval data for which statistical tests involving difference of means and difference of proportions is needed, check for inherent assumptions of the statistical models that are included in this course, interpret relationships in bivariate data, compare distributions of responses, understand the role of statistics in analyzing data and in inference.

In addition, students will have an opportunity to work on a project to obtain needed information for the college. Students will work together to: design and pilot an instrument for gathering data, obtain and collect data from a random sample of Oxford College students, evaluate data, prepare a final report that presents the class findings, present the report to the appropriate audience.

Grading: Grades will be determined by student performance/participation on four different problem sets, one individual experiment, the class project, and a final exam. Each problem set will have 125 possible points; the experiment will have up to 50 points; the class experiment will have a potential 200 points; the final exam will have 250 points, for a total of 1000 possible points:

4 problem sets @ 125	500
1 experiment	50
1 Class Project	200
1 final exam	250

Total	1000 points

There will be opportunities for extra credit work. Extra credit problems are usually not the standard type problems but those that require more thought and some original work. **You must do work independently for credit on bonus work unless otherwise indicated.** Points earned on extra credit will be added to your total points that will determine your grade.

In general,

900 points and up	A, A-
750 to 899 points	B+, B, B-
600 to 749 points	C+, C, C-
below 600 points	F

Some Policies:

All problem sets will be given out at least one week before the due date. All problem sets are due at class time on the dates indicated. Fifty points will be deducted per day for late problem sets. Emergencies will be handled on an individual basis.

For work on problem sets, you may use your own notes (those you have taken in class), the notebook provided for this class, your textbook, the computer facilities, and/or your own computer or calculator; however, you may not receive help from another person or talk to anyone about the problems on the problem sets.

Policies regarding your individual experiment will be explained in class, on the day the experiment is assigned.

The class project will be in the form of a survey, will be designed by the class, and will be related to course evaluations. We will attempt to answer, "What questions would students like asked on the college course evaluation sheet so that responses will include what students consider important for evaluation?"

Homework problems (assignments attached) may be worked with other members of this class. Solutions to some homework problems are on reserve at the library or in the notebook for this class. You will need to keep your class notes and homework problems well organized and complete so that they will be useful to you on your problem sets and the final exam. Class attendance is important. You are responsible for work done in class. There are no tutors for this course.

There is a Math 207 class conference on LearnLink. Please use the conference to check announcements, to communicate concerns, to pose questions appropriate for the class, and to attach work related to the class experiment.

Responsibilities**:**

* Each **student** has the following responsibilities:

1. Come prepared and on time to every class.
2. Complete all work on time with proper thought.
3. Consider that it is not always the fault of the instructor if the student doesn't understand the material. Use your outside help (office hours, SI sessions, eReserves)
4. Treat the instructor and peers with respect.
5. Ask questions. Asking questions is a sign of maturity, not ignorance, as long as the student thinks clearly before asking.
6. Understand that the instructor is not trying to "nit pick" when grading and remember that grading is the responsibility of the instructor. Accuracy is important in this class!

***** The **instructor** has the following responsibilities:

1. Come prepared to every class.
2. Design each class so students can accomplish the cognitive objectives listed in the syllabus.
3. Provide appropriate tips for studying and study materials as seem appropriate.
4. Create a mutually respectful classroom environment.
5. Return tests and quizzes in a timely manner so that students will know their grade.
6. Grading, as far as possible, to be consistent and impersonal even though students might not agree with the decisions concerning partial credit.

THE HONOR CODE OF OXFORD COLLEGE APPLIES TO ALL WORK IN THIS CLASS! YOUR PROBLEM SETS AND ANY OTHER WORK ARE PLEDGED TO BE YOUR WORK IN ACCORDANCE WITH INSTRUCTIONS GIVEN FOR THE ASSIGNMENT.

Important dates:

Friday, February 9	Individual Experiment due
Wednesday, February 14	Problem Set I due
Friday, March 9	Problem Set II due
March 12-16	SPRING BREAK
Friday, April 6	No Formal Class, work on the Class Project
Wednesday, April 11	Problem Set III due
Friday, April 20	Class Project due
Friday, April 27	Problem Set IV due
Monday, April 30	Last class day
	Final exam according to the college schedule

Homework Assignments

Wednesday, January 17 — Introduction
Friday, January 19

Read Chapter 1 and part of Chapter 2 (pages 4-26) to get an overview
Read Introduction and Chapter 1 in *Statistics You Can't Trust*.

PROBABILITY

Monday, January 22 — Definitions, Addition Rule, Multiplication Rule
Wednesday, January 24 — (conditional), Bayes Theorem, Jerome Cardan,
Friday, January 26 — Discuss Class Project

p. 45-46: 1, 2, 3, 5-27
Definitions: experiment, sample space, certain event, impossible event, mutually exclusive, independent

Monday, January 29 — Birthday Problem, Craps,
Wednesday, January 31

Individual Experiment assigned Wednesday
Problems on probability in the notebook.
Read Chapter 9 in Statistics You Can't Trust.

1. Who was Chevalier de Mere?
2. Why were Pascal and Fermat credited with the beginnings of probability?
3. Explain why "before the fact" and "after the fact" gives us different prospectives on determining probability. Relate these approaches to coincidences.
4. What is the classical approach?

COUNTING

Friday, February 2 — Fundamental Theorem of Counting, permutations, combinations

P. 48: 28-37, 40 (Get common denominator)

Monday, February 5 — Probability and Counting, Discuss Class Experiment (and any time during class as needed from now on…)

Problems on Counting in the notebook,
P. 50: 71-80, 82
Receive Problem Set I

FUNCTIONS (THEORETICAL)

Wednesday, February 7 Random Variable, Density Functions

Read in Chapter 2 pages 27 to 45; page 48: 41-45
Explain the difference between a pdf and a cdf for continuous and discrete functions.

Friday, February 9 Joint Density Functions, Marginal Distributions
Monday, February 12

page 49: 47-52, 58, 60, 62, 64, 68; page 52: 91, 92, 93, 97
Notebook problems
Know the following definitions: random variable, joint probability distribution, marginal distribution, conditional distribution, independence.
Individual Experiment Due on Friday

DISCRETE FUNCTIONS

Wednesday, February 14 Discrete Functions, Binomial Distribution

Read pages 53-72
page 92: 1, 2, 5, 6, 7, 9, 12, 13, 15, 16, 17, 18, 20
Problem Set I due

Friday, February 16 Poisson, Hypergeometric, and Uniform Distributions

page 94: 22, 24, 25, 29, 33, 34
problems in notebook
Know the following definitions: expectation, moments (about the origin and about the mean), mean, variance, skewness, kurtosis.

CONTINUOUS FUNCTIONS

Monday, February 19 Continuous Functions, Chebyshev's Theorem,
Wednesday, February 21 Normal Distribution, empirical rule, Central Limit Theorem

page 95: 37, 42, 44
page 96: 46-50, 54, 55, 57, 58, 59, 62, 64
page 100: 101, 103, 104, 105, 106, 108

Friday, February 23 Normal to approximate the binomial as a model, uniform

page 97: 66, 67, 68
page 99: 83, 85, 87, 94, 95
notebook problems

Monday, February 27 Review

Receive Problem Set II

DATA AND MEASUREMENTS

Wednesday, February 28 Descriptive Statistics
Friday, March 2 Class Project

Read pages 102-109; page 117: 1-5
Read Chapters 2, 3, and 10 in Statistics You Can't Trust.
Chapter 2
1. What are the two meanings of the word, statistics?
2. Why is an operational definition important?
3. Give one example of a meaningless statistic.
4. Give one example of a hyperaccuracy.
5. Define a self-selected sample. Give an example.
6. Define an unknowable statistic. Give an example.

Chapter 3
1. What are dishonest charts? Find an example from a magazine or newspaper.
2. Is it possible, in your opinion, to have a visually accurate pictogram? Why or why not?

Chapter 10
1. Why should one be suspicious of sample selection?
2. What is the difference in induction and deduction? Relate to statistical inference.
3. What is the margin of error and why is it important to report this margin when giving statistical information?

Monday, March 5 Data measures and classification
Wednesday, March 7 Class Project

Read pages 129 to 134; p. 162: 10, 11
problems in notebook
Know the following definitions: random sample, stem-and-leaf, outlier, statistical inference, histogram, quartiles, parametric vs nonparametric statistics, types of data (nominal, interval, ordinal, ratio), measures of central tendency and of variability.

Read Chapters 4 and 5 in Statistics You Can't Trust.
Chapter 4
1. What is a pseudoaverage?
2. Under what circumstances would the average (arithmetic mean) not be a good measure of central tendency? (The median is robust to what?)

Chapter 5
1. Why is the variation important in a set of data?

CONFIDENCE INTERVALS

Friday, March 9 Confidence Intervals for means and for proportions

p. 164: 35, 36; p. 168: 72
Know the following terms: inferential statistics, point estimate, maximum error of estimate
Problem Set II due

Spring Break, March 13-17

HYPOTHESIS TESTING

Monday, March 19 Class Project
Wednesday, March 22 Theoretical hypothesis testing
Friday, March 23

p. 118: 13, 14, 16, 18, 19, 27, 28, 30
definitions in the notebook
Read chapters 6, 7, 8 in Statistics You Can't Trust.
Chapter 6

1. Give the definitions of the following: percent, percent change, and percent points of change. How can each be misleading?
2. What is the opportunistic construction of a percent?
3. Explain why broad-base fallacy is a statistical trick. Make up an example to illustrate this "unscrupulous" behavior.

Chapter 7

1. Would a clear operational definition eliminate problems with "an untidy comparison" or are there other considerations?
2. Describe the *risk-your-life-and-live-longer-fallacy*. Find an example in a newspaper or magazine.
3. Why is a control group necessary to arrive at conclusions, especially in health related experiments?

Chapter 8

1. Give an example why human inclination is to make order or over explain.
2. For each of the three claims on pages 150-152, offer a Plausible Alternative Conclusion.

Monday, March 26 Hypothesis testing - means [one sample, large and small sample)]

problems in the notebook

Wednesday, March 28 — Means continued [two samples (dependent, large
Friday, March 30 — sample, small sample, homogeneity of variance)]

Read pages 138 to 159; p. 162: 12, 15, 16, 25, 26, 27
problems in the notebook

Monday, April 2 — Hypothesis Testing - proportions

p. 163: 28, 29, 30, 32, 33; problems in the notebook

Wednesday, April 4 — Review hypothesis testing

p. 164: 35, 36, 41, 42, 43, 45, 46, 47, 48, 49, 53, 55, 56; p. 169: 80, 81, 83
Receive Problem Set III

Friday, April 6 — No formal class, work on class project

REGRESSION

Monday, April 9 — Correlation and Simple Linear Regression

Read chapter 7
page 211: 2, 3, 7, 9, 13, 19
problems in the notebook
Know the following definitions: bivariate data, coefficient of determination, covariance, method of least squares, spuriously correlated

Wednesday, April 11 — ANOVA/Status of Class Experiment Evaluated

Problem Set III due
Work problems in the notebook

NONPARAMETRIC STATISTICS

Friday, April 13 — Chi Square Tables, Multinomial Experiments, median test/Finish Class Experiment
Monday, April 15

page 266: 1, 2, 3, 4, 5, 15, 18
problems in the notebook
Read Chapter 11, 12 in Statistics You Can't Trust.
Chapter 11

1. Why is it not appropriate to discuss cause and effect in relation to a regression? What is the difference between correlation and causation?
2. Find an example of *post hoc ergo propter hoc* fallacy (or *post hoc* fallacy) in a magazine or newspaper.
3. Explain the following terms: necessary cause, sufficient cause, necessary and sufficient cause, contributory cause.
4. What is a lurking variable and hwy should one not discount it?
5. What are some problems with the use of a small sample size?

Chapter 12

1. What does it mean to be “trapped in a cell a”?

Wednesday, April 18 — Wilcoxon Rank-Sum, Kruskal-Wallis
Friday, April 20

Work problems in notebook
Class Project due - Friday
Receive Problem Set IV on Wednesday

Monday, April 23 — Discuss Major Misuses of Statistics,
Wednesday, April 25 — Evaluate, Present Class Project, Discuss the
Friday, April 27 — Final Exam
Monday, April 30

Read Chapters 13 and 14 in Statistics You Can't Trust.
Chapter 13

1. Why is it easy to misinterpret conditional probabilities?
2. Explain the regression fallacy proposed by Sir Francis Galton.
3. Why should one be cautious of: (a) projections? (b) computer results?
   (c) omission of details? (d) ill-conceived ratios.

Chapter 14

1. Evaluate the four questions in bold print on page 241. Do these four questions include all the types of misinformation?
2. Pick your favorite three examples in this chapter. Give the number of your example and explain why it was selected by you.

finish notebook
Read Glossary in Statistics You Can't Trust
Problem Set IV due Friday
Receive Take Home part of final exam on Monday
Final Exam according to college schedule