Math 207, Probability and Statistics
Spring Semester, 2005
Dr. Evelyn Bailey

Office hours: Room 303 Seney Hall,
In general, 10 - 11:30 TTh; 2:00 – 3:30 MW, and others by appointment

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), 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, multinomial, Wilcoxon Rank Sum Test, Kruskal-Wallis Test).

Goals: At the end of this course students 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.

Grading: Grades will be determined by student performance on four different problem sets, one individual experiment, a class experiment, and a final exam. Each problem set will have 150 possible points; the experiment will have up to 60 points; the class experiment will have a potential 120 points; the final exam will have 220 points, for a total of 1000 possible points:

4 problem sets @ 150	600
1 experiment	60
1 Class Experiment	120
1 final exam	220

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 experiment will be in the form of a survey, will be designed by the class, and will emphasize: What students think about the honor system (Honor Code, Honor Council)?

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, and to pose questions appropriate for the class, and to attach work related to the class experiment.

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:

Wednesday, February 16	Problem Set I due
Wednesday, March 9	Problem Set II due
March 14-18	SPRING BREAK
Friday, March 25	No Class
Friday, April 8	No Class
Wednesday, April 13	Problem Set III due
Monday, April 18	Class Experiment due
Friday, April 29	Problem Set IV due
Monday, May 4	Last class day; Individual Experiment due
	Final exam according to the college schedule

Math 207

Homework Assignments

Wednesday, January 19 Introduction
Friday, January 21

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

PROBABILITY

Monday, January 24 Definitions, Addition Rule, Multiplication Rule
Wednesday, January 26 (conditional), Bayes Theorem, Discuss Class Experiment
Friday, January 28

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

Monday, January 31 Jerome Cardan, Birthday Problem, Craps,
Wednesday, February 2

Problems on probability in the notebook, pgs. 4,5
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 4 Fundamental Theorem of Counting, permutations, combinations

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

Monday, February 7 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 9 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 11 Joint Density Functions, Marginal Distributions
Monday, February 14

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.

DISCRETE FUNCTIONS

Wednesday, February 16 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 18 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 21 Continuous Functions, Chebyshev's Theorem, Normal
Wednesday, February 23 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 25 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 28 Review

Receive Problem Set II

DATA AND MEASUREMENTS

Wednesday, March 2 Descriptive Statistics
Friday, March 4

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 7 Data measures and classification
Wednesday, March 9

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?

Problem Set II due

CONFIDENCE INTERVALS

Friday, March 11 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

Spring Break *March 14-18*

HYPOTHESIS TESTING

Monday, March 21 Theoretical hypothesis testing

p. 118: 13, 14, 16, 18, 19, 27, 28, 30
definitions in the notebook

Wednesday, March 23 Hypothesis testing - means [one sample, large and small sample)]

problems in the notebook

Friday, March 25 NO class

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 28 Means continued [two samples (dependent, large
Wednesday, 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

Friday, April 1 Hypothesis Testing - proportions

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

Monday, 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

REGRESSION

Wednesday, April 6 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

Friday, April 8 NO class

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"?

Monday, April 11 ANOVA/Status of Class Experiment Evaluated
Wednesday, April 13

Problem Set III due Wednesday
Work problems in the notebook

NONPARAMETRIC STATISTICS

Friday, April 15 Chi Square Tables, Multinomial Experiments, median
Monday, April 18 test/Finish Class Experiment

Individual Experiment assigned Monday
page 266: 1, 2, 3, 4, 5, 15, 18
problems in the notebook

Wednesday, April 20 Wilcoxon Rank-Sum, Kruskal-Wallis
Friday, April 22

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

Monday, April 25 Discuss Major Misuses of Statistics, Catch Up
Wednesday, April 27

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.

Friday, April 29 Review, Evaluate
Monday, May 4

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