MAT 2376 – Applied Statistics

Winter 2008

Instructor: Dr. Brian Gill **Office:** OMH 209 **E-mail:** bgill@spu.edu
Phone: 281-2954 **Web page:** http://myhome.spu.edu/bgill

University and Departmental Mission: Seattle Pacific University seeks to be a premier Christian university fully committed to engaging the culture and changing the world by graduating people of competence and character, becoming people of wisdom, and modeling grace-filled community. The mathematics department at Seattle Pacific University seeks to provide excellent instruction to enable our students to be competent in the mathematics required for their chosen fields, and to share our expertise with the community through service and leadership. Hence, common goals for students in mathematics courses include 1) becoming competent in the topics covered in the course, 2) demonstrating skills and attitudes which contribute to professional, ethical behavior, 3) the ability to communicate mathematically, in both written and verbal form, and 4) learning to appreciate the beauty and utility of mathematics.

Course Overview: Statistics is the science of collecting and analyzing data for the purpose of drawing conclusions and making decisions. One of the primary goals of statistics is ***inferential statistics***, which can be defined as drawing conclusions and/or making decisions concerning a large population based only on data about a *sample* from the population. Formal methods of inferential statistics are based heavily on ***probability theory***, which was the subject of MAT 2375. Before studying inference, we will examine some further topics from probability, including more study of continuous probability distributions, moment generating functions, multivariate probability distributions, independent random variables, sampling distributions, and the central limit theorem.

For the remainder of the course, our primary goal will be to study two main topics from inferential statistics: ***estimation*** and ***hypothesis testing***. Even when data is carefully and properly collected, a sample will not typically mirror the population exactly, and different samples from the same population will each give somewhat different results (this is called *sampling variability*), so estimates obtained from samples will involve some sampling error. We will study some properties of good estimators and construct *confidence intervals*, which provide a probabilistic "margin of error" for our estimates. Finally, we will look at hypothesis testing, which involves making a decision between two competing claims (hypotheses) about the population based only on data about a sample.

Learning Objectives: By the end of the course, you should acquire:

- additional knowledge of probability theory and an understanding of the role that it plays in statistics;
- an understanding of multivariable probability distributions and independence;
- an understanding of fundamental statistical concepts, including randomization, estimation, confidence, testing, and significance;
- an understanding of fundamental concerns involved in data collection;
- an understanding of the central limit theorem and the role that it plays in statistics;
- knowledge of some basic properties of estimators, including bias, variance, and efficiency;
- the ability to perform a wide variety of statistical calculations and derivations;
- facility with approaching and solving practical problems and analyzing genuine data through statistical and probabilistic reasoning;
- the ability to use a computer to analyze data and to solve statistical problems from a variety of fields; and
- skills of communicating the results of statistical analyses through graphical and verbal means.

In addition to the specific content oriented objectives above, you should

- be able to prove basic theorems relating to the content of the course;
- be able to provide written explanations of the ideas behind key concepts from the course;
- have improved skills at problem solving and critical thinking;
- have improved skills at solving complex, multi-step problems; and
- have fun accomplishing these objectives, even if the material is difficult and takes a lot of time and effort.

Office Hours: My office hours will be announced in class during the first week of class and will be posted on the course web pages and outside my office door. You are strongly encouraged to drop by my office to ask questions, discuss problems, and just to get to know me better. If you are unable to meet with me during my scheduled office hours, I am available at other times by appointment. I also maintain an "open door" policy at my office – any time that my door is open you are welcome to drop in to talk to me, even if it is not during my scheduled office hours. *Please note that I also work as the statistician for a research lab at the University of Washington, so I am not available on campus on Tuesdays and Thursdays. Please plan accordingly.*

Prerequisites: This course is a direct continuation of MAT 2375 (Probability Theory), so you *must* successfully complete that course before enrolling in MAT 2376. Probability theory forms the foundation for the material in this course, so it is essential that you have a solid grasp of the content of MAT 2375. In addition, mastery of the material from MAT 1225 and 1226 (Calculus I and II) is required. Specifically, we will use techniques of differentiation and integration, improper integrals, integration by parts, L'Hopital's rule, and double integrals. While not absolutely essential, MAT 1228 (Series & Differential Equations) is also *strongly* recommended because infinite series will occasionally be used in this class.

Course Materials:

Textbook and Class Handouts: The primary materials for the course will be a collection of handouts given out in class. In addition, we will make use of selected materials from two textbooks: *Investigating Statistical Concepts, Applications, and Methods* by Beth Chance and Allan Rossman (these materials were available in the fall as a custom textbook through the SPU bookstore) and *Modern Mathematical Statistics with Applications* (1st edition) by Devore and Berk (the required chapters of this book will be available for you to purchase as an e-book for approximately $20. Details about purchasing the e-book will be provided in class.

Computer Software: We will make use of both Microsoft Excel and the statistical software package Minitab to perform calculations, to create graphics for analyzing data, and to conduct simulations to approximate long-run behavior of random phenomena. Minitab is available for your use on all computers in labs on campus. If you wish to use Minitab on your own computer, you can purchase a license at www.e-academy.com/minitab. The cost is $29.99 for a license which expires after 6 months, $49.99 for a 12-month license, or $99.99 for a permanent license.

Grading and Course Expectations

Attendance: Attendance will not be taken, but if you expect to succeed in this course, it is essential that you come to class every day. Unless you have an acceptable excuse ***and*** make special arrangements with me ***before*** class begins, missing an exam or failing to turn in an assignment on time will result in a grade of zero. Late homework will not be accepted for any reason, but if you have an acceptable excuse and contact me before class, I will drop the homework score.

NOTE: Things such as oversleeping, lack of preparation, or sneezing twice are NOT acceptable excuses. Acceptable excuses include a death in your immediate family or a *severe* illness, and *you are responsible for providing me with documentation of your excuse.*

Homework: The only way to truly learn mathematics is to work as many problems as possible. There will be homework assignments collected virtually every class period. Homework must be turned in by the time class starts on the day it is due; late homework will not be accepted for any reason. However, your lowest homework score will be dropped from your final grade.

You are strongly encouraged to come to my office to ask me questions about the homework. You are also encouraged to work with other students on the homework, but unless otherwise indicated for a particular assignment, you must individually write up and turn in your own solutions. You are also required to list on your paper any other individuals that you worked with or that gave you assistance with the homework – failure to do so will be considered cheating (turning in someone else's work as your own). Homework must be neat and easily readable or you will receive NO credit. You must show all of your work–a correct answer with no justification will also be worth NO credit.

Exams: There will be two exams: a midterm and a final. The midterm is tentatively scheduled for Friday, February 8 and will be worth 25% of your course grade. The cumulative final exam will be worth 35% of your overall grade and will be given in two parts. The first part will be take-home, and the second part will be in-class during the regularly scheduled final exam period (8:00-10:00 on Thursday, March 13). Most of the questions on the in-class exams will be *very* similar to exercises from the homework. You also may be expected to be able to state some definitions and to provide proofs of certain key theorems from the course.

Course Grades: Homework assignments will be worth 40% of your course grade, the midterm exam will be worth 25%, and the final exam will count for 35% of the overall grade. Course grades will be based on the following scale:

93-100% A		90-92% A-
87-89% B+	83-86% B	80-82% B-
77-79% C+	73-76% C	70-72% C-
67-69% D+	60-66% D	Below 60% E

A grade of I (incomplete) is only given for non-academic reasons such as a severe illness that prevents you from completing the course. You must have a passing grade on the material that you have completed in order to receive an incomplete.

Academic Dishonesty: Academic dishonesty includes copying another's work on an exam, preparing for an exam by using test questions from a stolen exam, bringing concealed answers to an exam, turning in another person's work as your own, committing plagiarism, or assisting another student in cheating. The *minimum* penalty for cheating or plagiarism in any form will be a zero for the assignment or exam in question. In addition, all students have an obligation to make efforts to prevent other students from cheating and to report incidents of cheating or plagiarism. Further details regarding SPU's academic dishonesty policies can be found online in the SPU Undergraduate Catalog. (http://www.spu.edu/acad/UGCatalog/20078/GeneralInfo/policies.asp).

Additional Notes:

Calculators: Calculators will be permitted on all exams. Any basic scientific calculator or a graphing calculator such as the TI-81, TI-82, or TI-85 will be allowed. However, calculators which are capable of symbolic manipulation (such as the TI-89 or TI-92) are *NOT* permitted. If you have any doubt as to whether or not your calculator is acceptable, please ask me as soon as possible.

E-mail: All SPU students have an SPU e-mail address. I will occasionally make use of these SPU e-mail addresses to send information to all members of the class, so you should check your e-mail regularly. If you do not use your SPU e-mail account, there is a utility available through Banner to set up your SPU e-mail account to forward messages to some other e-mail address. I strongly recommend doing this so that you do not miss any important messages.

Please note that while it can be a great tool for quick communication (such as scheduling an appointment to talk with me face-to-face), *e-mail is rarely a good substitute for face-to face conversations and is very poorly suited for answering mathematical questions*. When you come to my office to ask me questions, I engage you in a discussion about the problem, ask questions about what ideas you have for approaching the problem, explore various possible approaches (and what goes wrong with some of them), etc. In the process, I can usually find out precisely where your difficulties lie and help you to learn how to get past them. Such a conversation is impossible by e-mail. Furthermore, typing and e-mailing mathematical symbols is very time consuming, and the resulting equations in the e-mail e-mails often come out garbled (or even completely missing).

Students with Disabilities: Students with disabilities need to contact Disabled Student Services in the Center for Learning to request academic accommodations. Disabled Student Services sends letters out to all your professors indicating the appropriate accommodations for the classroom based on your disability. Once you have done this, you should also make an appointment to meet with me in person as early as possible in the quarter to discuss the details of how we will implement the accommodations in this course.

Tentative Class Schedule

The table below provides a very tentative schedule of topics to be covered in class. The exact dates on which material will be covered WILL vary somewhat from this list. Topics may be added to or removed from this list at any time.

Date	Topic
1/4	Continuous probability distributions (Ch.6 handout)
1/7	The gamma function and gamma distribution
1/9	Computer lab activity: Introduction to sampling distributions and the central limit theorem
1/11	More on continuous distributions
1/14	The moment generating function
1/16	Special continuous distributions: uniform, exponential, gamma, and normal
1/18	The normal distribution
1/21	**No Class:** Martin Luther King, Jr. Birthday
1/23	Wrap up of continuous distributions
1/25	Jointly distributed random variables
1/28	Independent random variables; Distributions of sums of independent random variables
1/30	Sampling distribution of the sample mean; The central limit theorem
2/1	More on the central limit theorem; Normal approximation to the binomial distribution
2/4	Introduction to estimation: The German tank problem; Properties of estimators: bias, variance, and efficiency
2/6	Pre-exam wrap-up/review
2/8	**Midterm Exam**
2/11	More on properties of estimators
2/13	Introduction to interval estimation, confidence, and margins of error
2/15	More on confidence intervals
2/18	**No Class:** Presidents' Day
2/20	Introduction to hypothesis testing for means
2/22	More on hypothesis testing
2/25	Estimating the population variance; the chi-square distribution
2/27	The chi square distribution and inference for a population variance
2/29	The *t*-distribution
3/3	Inference about the population mean
3/5	Inference about population proportions
3/7	Introduction to inference about the difference between two means
3/10	Wrap-up and review
3/13	**Final Exam 8:00-10:00**

Modifications to the course requirements can be made at any time. It is *your* responsibility to know all course requirements as described here or announced in class.