Stat 512
Statistical Inference

Syllabus
Autumn 2015

Instructor	:	Scott S. Emerson, M.D., Ph.D.			
		Professor of Biostatistics			
		Office	:	HSB H655-J	
		Phone	:	616-6678 (Biostatistics)	
		Email	:	semerson@uw.edu	
		Office hours	:	M	2:00 - 3:30
				F	1:30 - 2:30
				(or by appointment)	

Assistants	:	TA:	Nilanjana Laha (nlaha@uw.edu)
		TA:	David Whitney (davidw19@uw.edu)
		Grdr:	Hanzheng (Alexander) Zhai (zhaihz51@uw.edu)
		Office hours :	(see web pages for TA office hours)

Time and Place	:	Lectures	:	M	10:30a - 11:20a	HSB T747
				Tu	12:30p - 1:20p	HSB T733
				W	10:30a - 11:20a	HSB T733
				F	10:30a - 11:20a	HCK 320

Class Web Pages: http://www.emersonstatistics.com/s512

The web page will be used to post announcements, homework assignments, etc. I urge you to check this site regularly. Questions that are submitted to me (via email or otherwise) that I think might be of general interest will have their answers posted on the web page, as well.

Prerequisites : Stat 395 and Stat 421, 423, 504 or Biost 514,
or the equivalent (the latter courses can be taken concurrently),
or permission of the instructor

Text : Required: Casella and Berger (CB) *Statistical Inference, 2nd ed.*
Optional: Dudewicz and Mishra (DM) *Modern Mathematical Statistics*

Computing : Software : S-Plus or R

This is a class in statistical theory. However, some homework problems may not be able to be solved in closed form. In those cases, students may be required to perform numerical searches or numerical integration to obtain the answers.

Assignments : Written problem sets approximately weekly

Homework problems requiring a written solution are assigned approximately one week in advance of their due date. Homework assignments will be posted on the class webpages. We reserve the right to grade only selected portions of the written homework.

Students will learn the material best if they attempt to work all homework problems on their own: much is learned from exploring dead ends that did not work. However, after attempting solutions on their own, students are encouraged to seek help from the instructor, the TAs, and other students in order that they learn solutions that can be handed in.

However, the work that is handed in should reflect only that student's work. That is, obtaining help from other students in order to learn the METHODS of solution is allowed, but copying another student's answer is NOT ALLOWED and will be considered plagiarism. Thus, when preparing their final homework papers, students should reproduce their solution without consulting any notes, etc. derived from other students or the internet. (See comments about quizzes below.)

Problems handed in one day late will be discounted 50%. Problems handed in two days late will be discounted 75%. Students will not receive credit for problems handed in more than two days late except under the most extenuating of circumstances.

Homework Review Sessions:

On each Friday following the due date for a homework assignment, one or both TAs will review the solution to the assigned problems. Attendance at this session is entirely optional. The location of these sessions will be posted on the class web page.

Quizzes : Approximately every two weeks.

A short in-class, closed book, close notes quiz will be given during the first part of class approximately 5 times during the quarter. The day of the quiz will be announced at least 24 hours in advance. Problems on these quizzes will be taken exclusively from prior homework problems.

Exams : Midterm and final examinations will be in-class, closed book, closed notes.

Grading :

Written homeworks	25%
Quizzes	25%
Midterm	25%
Final examination	25%

Miscellanea :

1. Homework assignments will be posted on the course web pages approximately one week in advance.
2. Electronic mail (e-mail) will be used for communication of errata and other announcements that are of interest to the general class. It is the student's responsibility to ensure that his/her email is on the class email list as generated by the UW. Throughout the quarter, students may submit questions regarding the course material via e-mail. Answers to questions that I feel are of general interest will be broadcast to the entire class (the identity of the source of the question will be protected). Questions that are likely to be of interest only to a single student will usually be answered individually. I try for reasonably quick turnaround on email questions, but backlogs do occur. It may happen that I think I have answered your question in a general message broadcast to the class, but you are still unsure of the answer. Do not hesitate to send your question again, and I will try to address it further.
3. We regard academic integrity extremely important. Cheating and plagiarism will be fully investigated. Students are urged to familiarize themselves with UW policy:
   https://depts.washington.edu/grading/pdf/AcademicResponsibility.pdf

Course Content

This two quarter course sequence covers the mathematical theory behind standard statistical inferential techniques. It is targeted to graduate students majoring in statistics, biostatistics, and other disciplines requiring an understanding of statistical theory. The course starts with a review of the probability theory that is the basis for that inference. We will then cover both frequentist and Bayesian methods of point estimation, interval estimation, and hypothesis testing.

We will generally follow the organization of Casella and Berger, except that large sample optimality and results present in chapter 10 will typically be covered in parallel with the small sample results. At all time, greatest emphasis will be placed on those results that tend to be of greatest current use in applied methods and current development of new methods. I will try to indicate those results that are of great importance owing to their widespread applicability in the use of statistics to answer scientific questions, and those results that are perhaps of greatest interest owing to the possibility that the techniques of method development might generalize to further such development.

The first four chapters deal with probability theory, most of which should have been covered in prerequisites for this course. Nonetheless, I will cover material from these chapters as follows.

1. It is presumed that students are completely familiar with the content of chapter 1, and this material will be reviewed cursorily.
2. It is presumed that all students are familiar with the definition of random variables and common probability distributions. Coverage of chapters 2, 3, and 4 will primarily focus on a review of the most important probability distributions used in this course. Emphasis will be placed on the results that are most pertinent to scientifically justifiable data distributions, and distributions that are particularly useful in statistical inference.
3. The material of chapter 5 will be covered in some detail, especially as it pertains to large sample results in section 5.5.

Following the review of probability theory, we will cover frequentist methods of point estimation and optimality of estimators in chapters 6 and 7. This material will likely fill the remainder of STAT 512, with STAT 513 possibly continuing development of estimation methods and then covering material on theory of frequentist testing (both parametric and nonparametric), Bayesian methods, and decision theory.

At the end of Stat 512 - 513, you should be able to:

1. Derive point estimators in common parametric or strongly semi-parametric models using method of moments, maximum likelihood, least squares, and other common estimating equations.
2. Derive point estimators in a distribution free setting using method of moments and least squares estimating equations.
3. Derive exact probability distributions for estimators easily obtained in parametric problems.
4. Derive approximate large sample probability distributions for estimators in both parametric and distribution free problems.
5. Derive test statistics useful in both parametric and distribution free settings, along with their probability distributions.
6. Derive confidence intervals useful in both parametric and distribution free settings, along with their probability distributions.
7. Develop methods for estimating probability distributions for statistics based on both analytic and resampling methods.
8. Derive optimality criteria satisfied by those estimators and test statistics according to both frequentist and Bayesian criteria.
9. Derive the extent to which the inferential behavior of estimators and test statistics might be robust to departures from the assumptions under which they were derived.