THE UNIVERSITY OF UTAH

MDCRC 6000 Introduction to Biostatistics
Fall Semester 2018
Tuesday, 11:00am-12:50pm, HSEB 2912

Revision date: 8 August 2018

Instructor: Greg Stoddard (professional biostatistician)
Co-Director, Study Design and Biostatistics Center, http://medicine.utah.edu/ccts/sdbc/
Adjunct Assistant Professor, U of Utah School of Medicine, Department of Orthopaedics
Adjunct Associate Professor, U of Utah School of Medicine, Department of Family and Preventive Medicine
Home Department: Division of Clinical Epidemiology, Department of Internal Medicine
Email: greg.stoddard@hsc.utah.edu
Phone Number: 801-213-3774
Office Hours: By appointment only
Office Location: Williams Building, 295 Chipeta Way Room 1N433

Prerequisites: none
Semester Credits: 2

Continuing Education Option: MDCRC 600 is the continuing education course number of the class (non-credit) which meets with MDCRC 6000. If credit is not needed, this is a less expensive way to attend the class. University policy prohibits instructors allowing someone to simply sit in without registering.

Registration: The class is shown on the U of U Summer 2018 class schedule, under the department MDCRC (which stands for Medicine Clinical Research Center). To obtain the registration permission code, you can contact:

Ruben Rocha: ruben.rocha@hsc.utah.edu

Online Option: If you prefer, you can take the class as an on-line course (see *Teaching and Learning Methods* section below).

Required Materials
No textbook is required. The course manual, which is an electronic textbook written by the instructor, is provided as a free download (downloading either in Canvas or in Box). It is also found at the following website, so that you can get updates in the future (although this is currently off-line while some copyright issues are being resolved):

http://medicine.utah.edu/ccts/sdbc/stoddard_textbook.php

For now, it is online as a zip file in Box for anyone with a University ID number (uNID). To get access, copy the following web address into your Web browser:

https://uofu.box.com/Textbook

It is also provided as a zip file under the Files link on the Canvas home page of this course.

In this class, we will cover Chapters 1-1 to 1-8, 2-1 to 2-10, 2-15, 2-19, 3-4, 3-7, part of 4-1, part of 5-5, and part of 5-7 of the course manual. You can see the entire contents of the course by looking at these chapters in the course manual.

Computer Statistical software
You will be given, without cost to you, the Stata version 15 statistical software, on a 35 simultaneous user network license. The license permits the software to be installed on as many computers as the network license manager, Greg Stoddard, chooses, but only 35 users are allowed to be running the software at exactly the same time. All lectures and homework assignments will use Stata. After the class ends, you are welcome to keep it installed on your computer, as long as you are still associated with the University of Utah (it is a University of Utah site license). The software will not expire, but new versions come out every two years (version 16 will come out June 2019). Enough new things are added with each new version that it is always worthwhile to upgrade to the next version when it becomes available.

You do not have to be expert with Stata--examples of how to use it for specific applications are shown in the course manual.

How to install Stata/IC-15 will be provided in the syllabus in Canvas after you enroll in the course.

Stata/IC, where IC stands for "intercooled", whatever that means, contains the entire software package, but it does not allow extremely large datasets or utilize multiple processors on your computer. That is, it is a cheaper version, but you will not notice the difference. Bring a laptop to class, if you have one, for in-class practice. If you don't have a laptop, you can install Stata on your desktop computer to complete the homework and then just watch in class.

Stata Reference Manuals
When you install Stata, you have full access to the electronic copies of the reference manuals. To see them, open Stata. Then click on *Help* on the Stata task bar at the top of the Stata window. Then click on *PDF documentation* on the drop-down menu.

Optional Supplemental Videos
The official Stata website provides a long list of its own lecture videos which you might interesting. To see these, copy the following web address into your internet browser and go to that page:

http://www.stata.com/links/video-tutorials

Seeking Internet Help

There exists a wealth of information about Stata on the Internet. In an search engine, like *Google*, type your question or keywords and add *Stata* as a keyword. You will almost always find an answer to your question.

Course Description

Traditional statistics classes are taught using a formula approach. This class is very different. It teaches statistics using a research problem approach—how to solve the problem with software and what you need to know to solve actual research problems. A conceptual, intuitive approach to statistics is used, illustrating the concepts with simulation. Statistics topics include: errors in research, randomization, bias, confounding, descriptive statistics and graphs, levels of measurement, choosing a statistic, logic of significance tests, comparison of two groups, outliers, confidence intervals, reporting marginal significance, robustness to normality and equal variance assumptions, multiple comparison adjustments, power analysis, correlation, noninferiority testing, linear regression, logistic regression, log-rank test, Kaplan-Meier graphs, Cox regression, and epidemiology measures (risk difference, risk ratio, rate ratio, odds ratio, hazard ratio). In actual research, datasets always require some work to get them ready for statistical analysis. This class also covers this data management step of a statistical analysis, where the data management is performed in the Stata statistical software. These Stata topics include: getting data into Stata, creating and recoding variables, data cleaning such as removing invalid scores, merging files, assigning labels to variables and values, and reshaping datasets. It also teaches some basic computer programming constructs, such as: operators, functions, looping, and if statements. Finally, it covers creating scientific graphs of the quality required for submission of manuscripts for publication.

Technical Level of the Course

This is an applied course, not theoretical. Instead of emphasizing mathematical formulas, it will emphasize how to get the computer to do the computations for you. The class will utilize computer simulation to illustrate points, but you will not be required to do these yourself. The course will cover many things that are not taught in other statistics courses, so there will be a lot of new information even if you have had several previous statistics courses. A large number of solutions to practical problems, which other courses and statistics textbooks omit, will be covered in this class. The course will challenge a lot of ideas that statistics textbooks, instructors, statisticians, and journal reviewers propose, but justification for the challenge will always be provided in a supported, understandable way.

Course Outcomes

At the completion of this course, you will: be able to understand statistics reported in research articles; be able to prepare the patient characteristics table and statistical methods section for your own research; perform the statistical data analysis of the many commonly encountered statistical problems using the Stata statistical software; have a good grasp of the underlying logic behind applying statistics to research problems; be able to prepare a dataset for statistical analysis and create publication quality graphs for publication; and be able to respond to journal reviewers on a large number of common reviewer requests. The goal of the course is to prepare you to analyze your own data, not to turn you into a computer programmer, so just enough computer programming in Stata is covered to prepare you do this.

Teaching and Learning Methods (on-line and in-class options)
Recorded lectures of the material will be available in Canvas, but these are not "live" lectures that include the in-class discussion. Assignments and exams do not include material from in-class discussion, however, so the lecture videos, or even just reading the course manual, covers the material you are responsible for in terms of your class grade. If you choose, you can use these to take the course fully on-line. If you prefer taking the class by attending in-class lectures, however, you will have the benefit of clarifications diagrammed on the whiteboard and in-class discussions to reinforce concepts, but really the in-class lectures are simply covering the highlights of what is in the course manual. If you are attending the in-class lectures, you still might find the lecture videos useful if it helps to hear the lecture twice, or use them to make up any lectures you had to miss. Many on-line students find that reading the course manual is sufficient, and so they do not bother viewing the lecture videos. In fact, the course manual is designed to not require the video lectures. You can use any combination of reading, attending lectures, or viewing lecture videos. That is, you can take the class fully on-line, fully in-class, or a combination of on-line and in-class.

How to access the videos will be explained in the syllabus in Canvas after you enroll in the course.

The course manual is designed to be a look-up reference text for when you need it professionally, so it contains more material than the key points covered in the lectures. You are not responsible for the fine points, just the key points. If you can do the homework problems and answer the questions on the exams, you know you will have understood the key points that you are responsible for. Nobody is tracking whether you are attending class, reading the chapters, or watching the lecture videos. All that is tracked is the homework problems completed and uploaded into Canvas and the midterm exams taken in Canvas. In past years, the on-line students have done as well as the in-class students when in comes to completing homework correctly, answering exam questions correctly, and achieving an A grade for the course.

University Policies
This class complies with University policies:

1. ***The Americans with Disabilities Act***. The University of Utah seeks to provide equal access to its programs, services, and activities for people with disabilities. If you will need accommodations in this class, reasonable prior notice needs to be given to the Center for Disability Services, 162 Olpin Union Building, (801) 581-5020. CDS will work with you and the instructor to make arrangements for accommodations. All written information in this course can be made available in an alternative format with prior notification to the Center for Disability Services.
2. ***Addressing Sexual Misconduct***. Title IX makes it clear that violence and harassment based on sex and gender (which Includes sexual orientation and gender identity/expression) is a civil rights offense subject to the same kinds of

accountability and the same kinds of support applied to offenses against other protected categories such as race, national origin, color, religion, age, status as a person with a disability, veteran's status or genetic information. If you or someone you know has been harassed or assaulted, you are encouraged to report it to the Title IX Coordinator in the Office of Equal Opportunity and Affirmative Action, 135 Park Building, 801-581-8365, or the Office of the Dean of Students, 270 Union Building, 801-581-7066. For support and confidential consultation, contact the Center for Student Wellness, 426 SSB, 801-581-7776. To report to the police, contact the Department of Public Safety, 801-585-2677(COPS).

More information about University polices can be obtained at:

a. Student Code: http://regulations.utah.edu/academics/6-400.php
b. Accommodation Policy (see Section Q): http://regulations.utah.edu/academics/6-100.php

Assignments

Homework will be assigned each class session and will be of two types:

1. *Required problems to turn in for grading*. These problems are found in Chapters 2-97, 3-97, 4-97, and 5-97. Your solutions will need to be uploaded to Canvas, and will count toward your class grade.

2. *Additional practice problems not turned in for grading*. These are optional, so do them only if you want additional practice. These problems are found in Chapters 1-98, 2-98, 3-98, and 5-98, with solutions found in Chapters 1-99, 2-99, 3-99, and 5-99. Fair less effort has gone into writing the x-98 chapters as has gone into the x-97 chapters, so the x-98 chapters are incomplete. For that reason, I do not recommend bothering with this additional practice.

Exams: there will be three midterm exams. These will be take-home exams, and you will have one week to complete each exam. The exam questions are to be answered in Canvas.

Turning in Assignments and Exams Late

Although a deadline is always given for each homework assignment and exam, these are merely suggestions to encourage you to stay caught up with the class. If you turn in an assignment or exam late, you will still get full credit. Everything must be turned in by the end of finals week. If you wait to do everything during finals week, though, you will not learn as much and it will be a miserable week. If you have a very extenuating circumstance and require more time than finals week, you are welcome to seek permission for more time by emailing the instructor.

Grading

Grades will be based two-thirds on the lab problems and one-third on the exams. Each exam will have about eight questions. Some of the questions will be conceptual, and some require computing statistics using Stata. You can miss a total of 6 exam questions during the semester, which is equal to 2 questions per exam, and still get an A, miss 9 questions during the semester, which is 3 questions per exam, and get a B, and so on. For the structured labs, you will need to get 90% correct to get an A, 80% for a B, 70% for a C, and 60% or a D.

Updating the Class as I Go

I plan to add to the course manual throughout the semester, as there are a lot of topics I want to add. This is why I do not make all of the assignments and exams available up-front inside Canvas; which if I did, would permit you to complete the class on-line faster than the course proceeds in-class. Some of the on-line students will find this irritating, but everyone will benefit if I improve the course, so that is the way it is going to be.

This means the course manual will be changing several times during the semester, which means you will have to download the new version from time to time. If you want to type notes in a chapter, save it to a new file name, such as adding "notes added" to the end of the Word document file name, or save it to another folder outside of the course manual, so you do not copy over it later and lose your notes.

Course Schedule

Tentative Agenda (Will Update at the End of Each In-class Lecture)
The pace of the class will depend on how much in-class discussion occurs, but this tentative agenda is fairly close to how it will go.

(If taking the class on-line, then read the course manual chapters and view the recorded lectures if the course manual chapter was not perfectly clear. If attending the in-class lectures, there is no need to view the recorded lecture, but reading the course manual chapters will certainly reinforce the lectures.)

Date	Lecture Materials and Assignments	Topics Covered
Tue Aug 21	■ install Stata software ■ view recorded lecture "Ch 1-1 Stata windows" (6 min) ■ read Ch 1-1 "install Stata & recover windows" page 4 only	- changing Stata windows layout
	■ view recorded lecture "Ch 3-4 disjunctive syllogism Olestra example" (22 min) ■ read Ch 3-3 "Hill's Causal Criteria", page 3 only (introduce the Cheskin study) ■ read Ch 3-4 "Logic & Errors", page 2, beginning with section Disjunctive Syllogism through page 4	- disjunctive syllogism - logical argument of a scientific paper - organization of a scientific paper

	■ view recorded lecture "Ch 3-7 randomization" (16 min) ■ read Ch 3-7 "randomize using excel", all 17 pages ■ complete Lab_Aug21: Ch 3-97: ch 3-7 prob 1 ■ optional practice homework Ch 3-98: ch 3-7 problem 1 (preparing a randomization list)	- randomization methods
Tue Aug 28	■ view recorded lecture "Ch 1-2 reading in Stata formatted files" (8 min) ■ read Ch 1-2 "getting data into Stata", pp.1-4	- reading Stata formatted files (*use* command) - changing directory (*cd* command) - directory list (*dir* command) - clearing Stata memory (*clear* command and *clear* option)
	■ view recorded lecture "Ch 1-2 syntax and do-files" (12 min) ■ read Ch 1-2 "getting data into Stata", pp.5-10	- scrolling in Results window - syntax of Stata commands - Stata's help facility - Stata manual pdf files - suggested do-file structure
	■ view recorded lecture "Ch 1-2 reading Excel and ASCII files" (9 min) ■ read Ch 1-2 "getting data into Stata", pp.10-18 ■ complete Lab_Aug28: Ch 1-97: ch 1-2 probs 1 and 2 ■ optional practice homework Ch 1-98: ch 1-2 problems 1-3	- importing ASCII data files - importing Excel data files - importing ASCII data files when file names are not on first row, without specifying cell range
	■ view recorded lecture "Ch 2-1 frequency tables" (13 min) ■ read Ch 2-1 "describing variables", pp.1-6 ■ complete Lab_Aug28: Ch 2-97: ch 2-1 prob 1	- frequency tables - missing data
	■ view recorded lecture "Ch 2-1 histograms" (10 min) ■ read Ch 2-1 "describing variables", pp.7-16 ■ complete Lab_Aug28: Ch 2-97: ch 2-1 probs 2-4	- histogram - kernal density plot - mirror plot
	■ view recorded lecture "Ch 2-1 boxplots" (9 min) ■ read Ch 2-1 "describing variables", pp.17-25	- box plot - box plot outlier identification - clustered box plot

	■ complete Lab_Aug28: Ch 2-97: ch 2-1 probs 5 to 7	
	■ view recorded lecture "Ch 2-1 descriptive stats" (8 min) ■ read Ch 2-1 "describing variables", pp.26-28	- descriptive statistics definitions
	■ view recorded lecture "Ch 2-1 normal distribution" (5 min) ■ read Ch 2-1 "describing variables", pp.29-31	- historical account of normal distribution - normal distribution fits particularly well to distributions of errors - normal distribution formula - Stata "twoway function" for graphing any function
	■ view recorded lecture "Ch 2-1 normal approximation to distribution" (7 min) ■ read Ch 2-1 "describing variables", pp.32-40	- normal approximation to histogram - mean and standard deviation properties of Normal distribution - purpose of mean±SD notion - left and right skewness
Tue Sep 4	■ view recorded lecture "Ch 1-3 drop gen replace inlist" (11 min) ■ read Ch 1-3 "cleaning data" pp.1-6 ■ complete Lab_Sep4: Ch 1-97: ch 1-3 probs 1 and 2	- data cleaning - drop command - generate (gen) command - replace command - inlist( ) function
	■ view recorded lecture "Ch 1-3 storage_types string_missing recode" (8 min) ■ read Ch 1-3 "cleaning data", pp.6-10	- cannot do arithmetic on string variables - describe command - variable storage types - string missing (null string) - full audit trial do-file structure - destring command - recode command
	■ view recorded lecture "Ch 1-3 upper destring_force" (6 min) ■ read Ch 1-3 "cleaning data", pp.10-14 ■ complete Lab_Sep4: Ch 1-97: ch 1-3 probs 3 to 6 ■ optional practice homework Ch 1-98: ch 1-3 problems 1-5	- upper and lower functions - "force" option on destring command
	■ view recorded lecture "Ch 2-1 standard error" (10 min) ■ read Ch 2-1 "describing variables", pp.41-44 ■ complete Lab_Sep4: Ch 2-97: ch 2-1 prob 8	- standard error of the mean (SEM) - verify using a Monte Carlo simulation Gauss' claim (on page 29 that measurement errors (errors in measuring the mean) follow a normal distribution

		- verify formula for the SEM with that same Monte Carlo simulation
	■ view recorded lecture "Ch 2-1 level of measurement" (21 min) ■ read Ch 2-1 "describing variables", pp.45-52	- levels of measurement - choosing the best descriptive statistic for a patient demographics table in manuscript - failure of mean±SD to bound middle 2/3 when distribution is skewed
	■ view recorded lecture "Ch 2-1 table and tabstat commands" (8 min) ■ read Ch 2-1 "describing variables", pp.53-57 ■ skip pp.58-75, which are specific to basic science) ■ complete Lab_Sep4: Ch 2-97: ch 2-1 probs 9 to 13	- "by" prefix to get subgroup analysis - table command for descriptive stats - tabstat command for descriptive stats
Tue Sep 11	■ view recorded lecture "Ch 1-4 append" (5 min) ■ read Ch 1-4 "merging files", pp.1-5	- appending files (same variables, different study subjects)
	■ view recorded lecture "Ch 1-4 one-to-one & match merge" (8 min) ■ read Ch 1-4 "merging files" pp.6-11	- match merging (different variables, same study subjects) - one-to-one merge (line by line merge) - match merging while matching on a study ID
	■ view recorded lecture "Ch 1-4 match merging with options" (9 min) ■ read Ch 1-4 "merging files", pp.12-21 ■ complete Lab_Sep11: Ch 1-97: ch 1-4 prob 1 ■ optional practice homework Ch 1-98: ch 1-4 problem 1	- match merging overwriting missing - match merging overwriting nonmissing - one-to-many match merging - dropping _merge variable when merging many files - *erase* command to delete files from hard drive
	■ view recorded lecture "Ch 2-2 statistical regularity" (15 min) ■ read Ch 2-2 "logic of significance tests", pp.1-10	- simple and conditional probability - probability distributions - laboratory reference ranges - statistical regularity (strong law of large numbers) - simulation of statistical regularity - sampling variability
	■ view recorded lecture "Ch 2-2 signal-to-noise ratio" (5 min) ■ read Ch 2-2 "logic of significance tests", pp.11-14	- signal-to-noise ratio simulation - signal-to-noise ratio form of all significance tests

	■ view recorded lecture "Ch 2-2 sampling distribution" (7 min) ■ read Ch 2-2 "logic of significance tests", pp.15-20	- sampling distribution of test statistic - t distribution
	■ view recorded lecture "Ch 2-2 definition of p value" (9 min) ■ read Ch 2-2 "logic of significance tests", pp.21-26 ■ complete Lab_Sep11: Ch 2-97: ch 2-2 prob 1	- definition of p value - desired probability is an untractable problem - proof by contradiction - demonstration by improbableness - dysjunctive syllogism - "acceptance region" is a misleading choice of words - same statistics work for basic science and population science (one size fits all)
	■ view recorded lecture "Ch 2-2 reporting SD vs SEM" (10 min) ■ read Ch 2-2 "logic of significance tests", pp.27-38	- when report SD and when to report SEM - 95% confidence intervals can overlap by 25% and still be statistically significant - cut-and-paste response to reviewer who tells you to report SD in place of SEM for study outcomes
	■ view recorded lecture "Ch 2-2 discrimination problem" (7 min) ■ read Ch 2-2 "logic of significance tests", pp.39-46 ■ skip pp.44-52, which are specific to basic science)	- discrimination problem - mean profile plot with error bars - bar graph with error bars
	■ view recorded lecture "Ch 2-3 choice of significance test" (7 min) ■ read Ch 2-3 "choice of significance test", pp.1-4, glance at rest of chapter to see what is there ■ optional practice homework Ch 2-98: ch 2-3 problems 1 to 2	- choosing a statistical test
	■ midterm 1 assigned (covers Aug 21 to Sep 11 material) – one week to complete	
Tue Sep 18	■ view recorded lecture "Ch 1-5 variable and value labels" (12 min) ■ read Ch 1-5 "labeling variables and values", pp.1-8 ■ complete Lab_Sep18: Ch 1-97: ch 1-5 probs 1 to 3	- labeling variables - labeling values of variables - temporarily not displaying labels with "nolabel" option - permanently removing labels

	■ optional practice homework Ch 1-98: ch 1-5 problem 1	
	■ view recorded lecture "Ch 2-4 test for dichotomous outcome" (19 min) ■ read Ch 2-4 "comparison of 2 independent groups", pp.1-10 ■ complete Lab_Sep18: Ch 2-97: ch 2-4 probs 1 and 2	- properties of a confounding variable - p values in Table 1 for clinical trials - asymptotic vs exact tests - independent and dependent variables - two proportions z test - chi-square test - two proportions z test and chi-square test are identical tests in the 2 × 2 crosstabulation table case - preview of Fisher's exact test
	■ view recorded lecture "Ch 2-4 Fishers exact test is conservative" (10 min) ■ read Ch 2-4 "comparison of 2 independent groups", pp.11-17	- why chi-square test with continuity correction should not be used (and so is not available in Stata - conservativeness of Fisher's exact test - "exact" test does not mean "better" test
	■ view recorded lecture "Ch 2-4 choosing between chi-square and Fisher" (10 min) ■ read Ch 2-4 "comparison of 2 independent groups", pp.18-23 ■ complete Lab_Sep18: Ch 2-97: ch 2-4 prob 3	- choosing between chi-square test and Fisher's exact test - minimum expected cell frequency rule-of-thumb - expected cell frequencies - Barnard's unconditional exact test
	■ view recorded lecture "Ch 2-4 Fisher-Freeman-Halton test" (6 min) ■ read Ch 2-4 "comparison of 2 independent groups", pp.24-26 ■ complete Lab_Sep18: Ch 2-97: ch 2-4 prob 4	- Fisher-Freeman-Halton test for $r \times c$ tables
	■ view recorded lecture "Ch 2-4 Wilcoxon-Mann-Whitney test" (11 min) ■ read Ch 2-4 "comparison of 2 independent groups", pp.27-32 ■ complete Lab_Sep18: Ch 2-97: ch 2-4 prob 5	- Wilcoxon-Mann-Whitney test
	■ view recorded lecture "Ch 2-4 WMW test porder option" (9 min) ■ read Ch 2-4 "comparison of 2 independent groups", pp.33-38	- Wilcoxon-Mann-Whitney test is not a comparison of medians - WMW test, P(X>Y) estimate of effect size
	■ view recorded lecture "Ch 2-4 WMW test another example and exact WMW test" (7 min) ■ read Ch 2-4 "comparison of 2 independent groups", pp.39-48	- example WMW test when medians are equal but p < .05 - reporting the WMW test P(X>Y) estimate of effect size - exact WMW test

		- rule of thumb for choosing between WMW test and exact WMW test
Tue Sep 25	■ view recorded lecture "Ch 1-6 graph editor" (5 min) ■ read Ch 1-6 "basic graphics", pp.1-3	- *graph display* to bring back the graphics window - graph editor - *graph describe* to see command that created the graph
	■ view recorded lecture "Ch 1-6 overlaid graphs" (6 min) ■ read Ch 1-6 "basic graphics", pp.4-7	- scatterplot (*twoway scatter* command) - linear fit plot (*twoway lfit* command) - smooth fit (*twoway fpfit* command) - side by side graphs (*by* option) - overlaid graphs
	■ view recorded lecture "Ch 1-6 profile plot" (7 min) ■ read Ch 1-6 "basic graphics", pp.7-10	- profile plot - *egen* command with *round* function to get means at each value of x-axis - first sort on x-axis variable before request line graph, or will get a strange choas graph - *#delimit ; #delimit cr* combination to continue command across several lines
	■ view recorded lecture "Ch 1-6 titles width scheme" (8 min) ■ read Ch 1-6 "basic graphics", pp.11-16	- graph features (titles, legends, tick marks) - line thickness - graph schemes
	■ view recorded lecture "Ch 1-6 s1color s1mono text" (4 min) ■ read Ch 1-6 "basic graphics", pp.17-22	- s1color and s1mono schemes - eliminating border around graph - turning off legend - text in graph
	■ view recorded lecture "Ch 1-6 color patterns error bars" (5 min) ■ read Ch 1-6 "basic graphics", pp.22-27	- white space between x-labels and x-title - line colors - line patterns - scatterplot with error bars - *rcap* graph (range line with caps to produce error bars)
	■ view recorded lecture "Ch 1-6 marker_size_color_thickness" (5 min) ■ read Ch 1-6 "basic graphics", pp.28-36	- xlabel option to get white space on left and right sides of graph - xlabels other than numbers - removing tick marks - symbol styles, colors, and thickness
	■ view recorded lecture "Ch 1-6 gee whiz graph log_scale" (5 min)	- white space top and bottom of graph - gee whiz graph

	■ read Ch 1-6 "basic graphics" pp.36-42	- use of logarithm scale for relative measures (odds ratio, relative risk, etc)
	■ view recorded lecture "Ch 1-6 special_characters label_sizes" (5 min) ■ read Ch 1-6 "basic graphics" pp.43-49	- special characters in text, superscripts, bold - modifying graph in PowerPoint
	■ view recorded lecture "Ch 1-6 submitting graphs to journals" (8 min) ■ read Ch 1-6 "basic graphics" pp.50-59 ■ complete Lab_Sep25: Ch 1-97: ch 1-6 probs 1 to 7 ■ optional practice homework Ch 1-98: ch 1-6 problem 1	- *.tiff and *.eps graphics files - RGB and CMYK color images - size of graph in inches
	■ view recorded lecture "Ch 2-4 central limit theorem" (11 min) ■ read Ch 2-4 "comparison of 2 independent groups", pp.49-53	- parametric and nonparametric tests - Central Limit Theorem (CLT) - simulation of CLT - CLT operates to make parametric tests robust to normality assumption
	■ view recorded lecture "Ch 2-4 two assumptions of t-test" (12 min) ■ read Ch 2-4 "comparison of 2 independent groups", pp.54-59 ■ complete Lab_Sep25: Ch 2-97: ch 2-4 prob 6 and 7	- two versions of independent sample t-test - homogeneity of variance (equal variance) assumption of independent sample t-test - Levene's equal variance test - normality assumption of t-test - Shapiro-Wilk normality test - robustness of t-test to these assumptions
	■ view recorded lecture "Ch 2-4 influence of outliers" (9 min) ■ read Ch 2-4 "comparison of 2 independent groups", pp.58-64	- influence of outliers on t-test - Wilcoxon-Mann-Whitney test is not influenced by outliers
	■ view recorded lecture "Ch 2-4 reporting styles" (16 min) ■ read Ch 2-4 "comparison of 2 independent groups", pp.65-71 ■ ----ended on page 70----	- truncation approach to outliers - how many decimals to report for descriptive statistics - how many decimals to report for p values - reporting styles for two group comparison of showing means or medians - protocol or statistical methods section suggestion for an interval scale variable comparison

Tue Oct 2	■ view recorded lecture "Ch 1-7 for loops" (13 min) ■ read Ch 1-7 "looping collapsing & reshaping", pp.1-8 ■ complete Lab_Oct2: Ch 1-97: ch 1-7 problems 1 and 2	- for loops (looping over numbers) - nested for loops - options to enhance *display* command - controlling logging inside do-file
	■ view recorded lecture "Ch 1-7 local_macros scalars" (3 min) ■ read Ch 1-7 "looping collapsing & reshaping" pp.8-11	- for loops over string constants - while loops - local macros - scalars
	■ view recorded lecture "Ch 1-7 egen reshape" (8 min) ■ read Ch 1-7 "looping collapsing & reshaping", pp.12-17 ■ complete Lab_Oct2: Ch 1-97: ch 1-7 problems 3 and 4	- egen functions: operations across rows or columns of data matrix - reshaping (pivoting) rows and columns
	■ view recorded lecture "Ch 1-7 tagging collapsing" (13 min) ■ read Ch 1-7 "looping collapsing & reshaping", pp.18-26 ■ optional practice problems: Ch 1-98: ch 1-7 problems 1 to 3	- tagging data rows - collapsing data into summary statistics - automated profile plot using collapsing
	■ ----begin on page 70---- ■ view recorded lecture "Ch 2-4 reporting styles" (16 min) ■ read Ch 2-4 "comparison of 2 independent groups", pp.65-71	- truncation approach to outliers - how many decimals to report for descriptive statistics - how many decimals to report for p values - reporting styles for two group comparison of showing means or medians - protocol or statistical methods section suggestion for an interval scale variable comparison
	■ view recorded lecture "Ch 2-4 Fisher-Pitman_test confidence_intervals" (23 min) ■ read Ch 2-4 "comparison of 2 independent groups", pp.72-84	- Fisher-Pitman permutation test for independent samples - confidence intervals - relationship of confidence interval and significance testing - statistical tests to identify outliers - discussing outliers in articles - pre-specification of analysis
	■ (optional) read Ch 2-4 "comparison of 2 independent groups", pp.85-92	- mathematric proof that two proportions z test is identically the chi-square test for 2×2 table

	■ view recorded lecture "Ch 2-19 robustness of t-test to assumptions" (24 min) ■ read Ch 2-19 "robustness of t-test", pp.1-33 basically skim read to see what is there	- robustness of t-test to normality and equal variance assumptions - Wilcoxon-Mann-Whitney is not a panacea, having problems of its own
Tue Oct 9	■ no class – Fall Break	
Tue Oct 16	■ view recorded lecture "Ch 1-8 operators" (11 min) ■ read Ch 1-8 "operators if dates and times", pp.1-6 ■ complete Lab_Oct16: Ch 1-97: ch 1-8 probs 1 and 2	- arithmetic operators - string operators - relational operators - logical operators - order of operations
	■ view recorded lecture "Ch 1-8 boolean arithmetic" (7 min) ■ read Ch 1-8 "operators if dates and times", pp.7-9	- Boolean arithmetic - truth tables for "and", "or", and "not"
	■ view recorded lecture "Ch 1-8 missing and greater than" (6 min) ■ read Ch 1-8 "operators if dates and times", pp.10-14	- missing data with arithmetic operators - missing data with relational operators - *cond( )*, condition function
	■ view recorded lecture "Ch 1-8 dates and times" (10 min) ■ read Ch 1-8 "operators if dates and times", pp.15-21 ■ complete Lab_Oct16: Ch 1-97: ch 1-8 probs 3 and 4 ■ optional practice problems: Ch 1-98: ch 1-8 problems 1 to 6	- working with dates and times - date literals (date constants) - generating a double precision time variable
	■ view recorded lecture "Ch 2-5 definition of power" (11 min) ■ read Ch 2-5 "basics of power analysis", pp.1-8 ■ complete Lab_Oct16: Ch 2-97: ch 2-5 problems 1 and 2	- sample size determination and statistical power are based on statistical regularity - monotonic increasing power with increasing sample size
	■ view recorded lecture "Ch 2-5 decision errors" (12 min) ■ read Ch 2-5 "basics of power analysis", pp.9-11	- decision errors (Type I and Type II errors) - diagnostic test logic - sensitivity and specificity - lack of statistical significance with high statistical power does not provide a logical argument for "no

		effect"
	■ view recorded lecture "Ch 2-5 one or two- sided comparisons" (13 min) ■ read Ch 2-5 "basics of power analysis", pp.12-18	- power is a function of 5 things - one- or two-sided comparisons (one- or two-tailed tests) - number of tails affects power - ethically bound to always report a two-sided comparison
	■ view recorded lecture "Ch 2-5 sample size and power calculation" (7 min) ■ read Ch 2-5 "basics of power analysis", pp.18-23 ■ complete Lab_Oct16: Ch 2-97: ch 2-5 problems 3 and 4	- choice of alpha affects power - difference in means affects power - SDs affects power - power calculation with *power* command - sample size calculation with *power* command
	■ midterm 2 assigned (covers Sep 18 to Oct 16 material) – one week to complete	
Tue Oct 23	■ view recorded lecture "Ch 4-1 chi-square fisher" (6 min) ■ read Ch 4-1 "specific applications", pp.7-11 ■ complete Lab_Oct23: Ch 4-97: ch 4-1 probs 1 to 3	- sample size and power for chi-square test - sample size and power for Fisher's exact test
	■ view recorded lecture "Ch 2-5 five items for power" (14 min) ■ read Ch 2-5 "basics of power analysis", pp.24-28	- 5 items required for power or sample size calculation - range/6 and range/4 rule for estimating SD - obtaining SD from IQR - starting with power and N, computing the smallest effect you can detect
	■ view recorded lecture "Ch 2-5 standardized effect size" (9 min) ■ read Ch 2-5 "basics of power analysis", pp.29-33 ■ complete Lab_Oct23: Ch 2-97: ch 2-5 problems 5 and 6 ■ optional practice problems: Ch 2-97: ch 2-5 problems 1 to 3	- standardized effects size (Cohen's d)
	■ view recorded lecture "Ch 2-5 popularity of standardized approach" (12 min) ■ read Ch 2-5 "basics of power analysis", pp.33-38	- reasoning behind Cohen's d - protocol suggestion using standardized approach - protocol suggestion using actual means and SDs - published examples of sample size paragraphs

	■ view recorded lecture "Ch 2-5 large sample p value problem" (4 min) ■ read Ch 2-5 "basics of power analysis", pp.39-40 ■ complete Lab_ Oct23: Ch 2-97: ch 2-5 problem 7	- large sample p value problem - replace p values with confidence intervals when sample sizes are very large
	■ remainder of Ch 2-5 requires some later chapters be covered first—we will return to the remaining pages later	< we will finish the chapter later in the course, after the relevant topics have been introduced >
	■ view recorded lecture "Ch 2-8 how multiplicity arises" (11 min) ■ read Ch 2-8 "multiplicity and comparison of 3 or more independent groups", pp.1-5	- multiple comparison problem - multiplicity (a fancier name) - Monte Carlo simulation of multiplicity - 5 ways multiplicity arises in clinical trials
	■ view recorded lecture "Ch 2-8 p value adjustment procedures" (14 min) ■ read Ch 2-8 "multiplicity and comparison of 3 or more independent groups", pp.6-10 ■ complete Lab_ Oct23: Ch 2-97: ch 2-8 prob 1	- formula for inflation of alpha - revised Monte Carlo simulation of multiplicity - multiple comparison adjustments are expected - p value adjustment procedures (Bonferroni and Holm's)
	■ view recorded lecture "Ch 2-8 propagation of Bonferroni" (7 min) ■ read Ch 2-8 "multiplicity and comparison of 3 or more independent groups", pp.11-14	- more p value adjustment procedures (Šidák, Holm-Šidák, Hochberg, Finner's, Hommel's) - unfortunate propagation of Bonferroni, resulting in loss of deserved research findings
	■ view recorded lecture "Ch 2-8 math proofs for Holms" (9 min) ■ read Ch 2-8 "multiplicity and comparison of 3 or more independent groups", pp.14-18	- mathematical proofs that Holm's maintains alpha as well as Bonferroni - protocol suggestion for describing multiple comparison procedure - correlated endpoints situation (Tukey-Ciminera-Heyse procedure)
	■ view recorded lecture "Ch 2-8 install mcpi" (5 min) ■ read Ch 2-8 "multiplicity and comparison of 3 or more independent groups", pp.18-20	- installing *mcpi* command and other commands in *ado files* section of course manual - *sysdir* command to reveal where Stata searches for *ado* files (program files)
	■ view recorded lecture "Ch 2-8 mcpi command" (7 min) ■ read Ch 2-8 "multiplicity and comparison of 3 or more independent groups", pp.21-24	- *mcpi* command to get adjusted *p* values

	■ complete Lab_Oct23: Ch 2-97: ch 2-8 prob 2	
	■ view recorded lecture "Ch 2-8 primary-secondary" (9 min) ■ read Ch 2-8 "multiplicity and comparison of 3 or more independent groups", pp.25-28	- primary-secondary approach to multiplicity
	■ view recorded lecture "Ch 2-8 misconception that anova is required (8 min) ■ read Ch 2-8 "multiplicity and comparison of 3 or more independent groups", pp.29-33	- misconception that analysis of variance (ANOVA) must precede pairwise multiple comparisons - conservativeness of ANOVA
	■ view recorded lecture "Ch 2-8 false discovery rate (6 min) ■ read Ch 2-8 "multiplicity and comparison of 3 or more independent groups", pp.33-38 ■ complete Lab_Oct23: Ch 2-97: ch 2-8 prob 3	- false discovery rate - Benjamini-Hochberg procedure
Tue Oct 30	■ view recorded lecture "Ch 2-8 win criteria approach (9 min) ■ read Ch 2-8 "multiplicity and comparison of 3 or more independent groups", pp.39-47	- win criteria (clinical decision rule) Approach - Moonblast energy drink example
	■ view recorded lecture "Ch 2-8 reporting comparisons separately approach (10 min) ■ read Ch 2-8 "multiplicity and comparison of 3 or more independent groups", pp.48-49	- reporting and discussing comparisons separately approach
	■ view recorded lecture "Ch 2-8 when to use global test (7 min) ■ read Ch 2-8 "multiplicity and comparison of 3 or more independent groups", pp.50-54	- global test only useful in Table 1 "patient characterisics" - comparing 3+ groups on dichotomous outcome - comparing 3+ groups on nominal outcome
	■ view recorded lecture "Ch 2-8 ordinal outcome comparison" (5 min) ■ read Ch 2-8 "multiplicity and comparison of 3 or more independent groups", pp.54-57	- comparing 3+ groups on ordinal outcome - Kruskal-Wallis analysis of variance

	■ view recorded lecture "Ch 2-8 interval outcome comparison" (6 min) ■ read Ch 2-8 "multiplicity and comparison of 3 or more independent groups", pp.58-64 ■ complete Lab_Oct30: Ch 2-97: ch 2-8 prob 4	- comparing 3+ groups on interval outcome - one-way analysis of variance
	■ view recorded lecture "Ch 2-13 confidence intervals" (11 min) ■ read Ch 2-13 "reporting CIs vs P values", pp.1-3	- confidence intervals contain much more information than p values - published example of using confidence interval graph (forest plot) to make claims without using p values
	■ view recorded lecture "Ch 2-13 marginally significant" (14 min) ■ read Ch 2-13 "reporting CIs vs P values", pp.4-12	- Fisher's retraction of p<0.05 cut-off for statistical significance - arguing 0.05<p<0.10 as marginally significant
	■ view recorded lecture "Ch 2-6 summing scale items" (6 min) ■ read Ch 2-6 "more on levels of measurement", pp.1-2	- sum of ordinal scale items produces an interval scale
	■ view recorded lecture "Ch 2-6 ordinal analyzed as interval" (12 min) ■ read Ch 2-6 "more on levels of measurement", pp.3-6	- visual analog scales (VAS) are interval scales - 11 tick marks is an interval scale - treating ordinal scale as interval opens you to criticism, but the criticism is not actually justified
	■ view recorded lecture "Ch 2-6 dichotomous is interval descriptive stats" (7 min) ■ read Ch 2-6 "dichotomous predictor variables", pp 6 to 10	- both outcome and predictor variables in linear regression model must have at least an interval scale - psychometrician argument that dichotomous variable (scored 0 or 1) is an interval scale - treating dichotomous variable as interval scale works for descriptive statistics demonstration - Bernoulli variable - population standard deviation
	■ view recorded lecture "Ch 2-6 dichotomous is interval test stats" (4 min) ■ read Ch 5-3 "dichotomous predictor variables", pp 11 to 14	- treating dichotomous variable as interval scale works for test statistics demonstration - t-test and chi-square test in 2 x 2 crosstabulation table are algebraically identical tests

Tue Nov 6	■ view recorded lecture "Ch 3-4 disjunctive_syllogism errors_in_publications" (15 min) ■ read Ch 3-4 "Logic and Errors", pp 1-7	- Hume's refutation of induction - Popper's deductive testing of theories - Proof in science - disjunctive syllogism - IMRAD structure of research article - Errors in published research
	■ view recorded lecture "Ch 3-4 paradigm_eras" (12 min) ■ read Ch 3-4 "Logic and Errors", pp 8-14	- Stoddard's aphorism - paradigm eras - prepostivist era - postivist era - postpositivist era - paradigm transitions - bloodletting and beliefs
	■ In-class analysis of bloodletting dataset (on-line students not responsible for this material—example of Ch 3-4 material)	
	■ In-class discussion of Cheskin paper (on-line students not responsible for this material—example of Ch 3-4 material)	
	■ view recorded lecture "Ch 2-10 extending_t-test" (8 min) ■ read Ch 2-10 "linear regression", pp.1-5 ■ complete Lab_Nov6: Ch 2-97: ch 2-10 probs 1 and 2	- mimicking bench experiment with regression model - linear regression extends t-test to allow control for potential confounders (covariates)
	■ view recorded lecture "Ch 2-10 controlling_for_covariates" (10 min) ■ read Ch 2-10 "linear regression", pp.6-11	- adding covariates to model - 10% change-in-effect variable selection rule - adding restriction to model
	■ view recorded lecture "Ch 2-10 mean-centering" (11 min) ■ read Ch 2-10 "linear regression", pp.12-15	- restriction vs representativeness - pragmatic (health services) research - mean-centered models
	■ view recorded lecture "Ch 2-10 holding_constant" (4 min) ■ read Ch 2-10 "linear regression", pp.15-21	- *predict* command to obtain predicted values - multivariable prediction equation
	■ view recorded lecture "Ch 2-10 perpendicular distance" (5 min) ■ read Ch 2-10 "linear regression", pp.21-28	- geometric intrepretation of holding constant - perpendicular distances between groups
	■ view recorded lecture "Ch 2-10 correlation" (9 min) ■ read Ch 2-10 "linear regression", pp.29-34	- correlation - Pearson r, Multiple R - coefficient of determination

	■ complete Lab_Nov6: Ch 2-97: ch 2-10 probs 3 and 4	
Tue Nov 13	■ read Ch 2-9 "correlation", pp.2-4 (no video lecture available)	- interpreting the size of correlation - futility of a rule of thumb for size of correlation
	■ read Ch 2-9 "correlation", pp.5-7 (no video lecture available)	- futility of correlation with small sample sizes
	■ view recorded lecture "Ch 2-5 overfitting" (9 min) ■ read Ch 2-5 "basics of power analysis", pp.41-47 ■ complete Lab_Nov13: Ch 2-97: ch 2-5 problems 8 to 9	- overfitting - graphical illustration of overfitting - rule-of-thumb to avoid overfitting
	■ view recorded lecture "Ch 5-5 risk study_designs" (10 min) ■ read Ch 5-5 "deriving logistic regression model", pp 1-6	- turning off variable label after generating a variable using the *recode* command - risk (incidence proportion) - prevalence (prevalence proportion) - cohort, cross-sectional, and case-control study designs - risk estimate not obtainable from case-control study
	■ view recorded lecture "Ch 5-5 epitab citicism_of_linear_regression" (9 min) ■ read Ch 5-5 "deriving logistic regression model", pp 7-10 ■ complete Lab_Nov13: Ch 5-97: ch 5-5 prob 1	- epitab (Tables for Epidemiologists) - risk difference with confidence interval - epitab and logistic regression require dichotomous variables be scored as 0-1 - linear regression of dichotomous outcome is a risk difference model - criticism of using linear regression with dichotomous outcome - risk difference model cannot be used with a case-control study design
	■ view recorded lecture "Ch 5-5 predicted probabilities" (5 min) ■ read Ch 5-5 "deriving logistic regression model", pp 11-14 ■ complete Lab_Nov13: Ch 5-97: ch 5-5 prob 2	- predicted probabilities from risk difference model - example of a dataset with predicted probabilities all within the 0 to 1 range - example of a dataset with a few predicted probabilities outside of the 0 to 1 range
	■ view recorded lecture "Ch 5-5 logistic function" (9 min) ■ read Ch 5-5 "deriving logistic regression model", pp 15-21	- logistic function - risk defined as logistic function consistent with dose-threshold response commonly observed in

		exposure-disease relationships - logistic regression predicts risk constrained to be between 0 and 1
	■ view recorded lecture "Ch 5-5 verifying prediction equation" (4 min) ■ read Ch 5-5 "deriving logistic regression model", pp 21-25 ■ complete Lab_Nov13: Ch 5-97: ch 5-5 prob 3	- predicted probabilities or linear predictor - manually plugging linear predictor predicted values into logistic function to obtain predicted probabilities - graphical verification that logistic regression curve was closer to the scatterplot data than the linear regression curve
	■ view recorded lecture "Ch 5-5 logit odds_ratio" (12 min) ■ read Ch 5-5 "deriving logistic regression model", pp 26-31	- definition of odds - logit form of logistic regression - odds ratio (OR) same for all three study designs - epitab to compute OR with confidence interval
	■ view recorded lecture "Ch 5-5 interpreting_odds_ratio" (6 min) ■ read Ch 5-5 "deriving logistic regression model", pp 32-35 ■ complete Lab_Nov13: Ch 5-97: ch 5-5 prob 4	- interpreting constant in logistic regression - interpreting odds ratio in logistic regression
	■ view recorded lecture "Ch 5-5 assumed_form_subgroup_analysis" (8 min)■ read Ch 5-5 "deriving logistic regression model", pp 36-43 ■ complete Lab_Nov13: Ch 5-97: ch 5-5 probs 5 and 6	- linear in log odds - exponential in odds ratio - percent change in odds interpretation - subgroup analysis
	■ view recorded lecture "Ch 5-5 adjusted_proportions" (5 min)■ read Ch 5-5 "deriving logistic regression model", pp 44-47 ■ complete Lab_Nov13: Ch 5-97: ch 5-5 prob 7	- adjusted proportions
Tue Nov 20	■ view recorded lecture "Ch 5-7 risk_ratio rate_ratio" (12 min) ■ read Ch 5-7 "Intro Cox regression & Kaplan-Meier Plots", pp 1-5	- life table data format - risk ratio definition and analysis - rate ratio definition and analysis
	■ view recorded lecture "Ch 5-7 hazard_ratio" (7 min) ■ read Ch 5-7 "Intro Cox regression & Kaplan-Meier Plots", pp 6-10	- binomial probability mid-p exact test for person-time data - hazard ratio analysis

	■ complete Lab_Nov20: Ch 5-97: ch 5-7 prob 1	
	■ view recorded lecture "Ch 5-7 goofing around analysis" (9 min) ■ read Ch 5-7 "Intro Cox regression & Kaplan-Meier Plots", pp 11-20	- form of survival analysis datasets - failure of goofing around analysis (risk ratio, rate ratio, t-test of follow-up time) fails
	■ view recorded lecture "Ch 5-7 life tables" (7 min) ■ read Ch 5-7 "Intro Cox regression & Kaplan-Meier Plots", pp 21-25 ■ complete Lab_Nov20: Ch 5-97: ch 5-7 probs 2 and 3	- life tables - actuarial estimates of hazard (risk) - Kaplan-Meier assumed estimates of hazard (risk)
	■ view recorded lecture "Ch 5-7 log-rank test" (8 min) ■ read Ch 5-7 "Intro Cox regression & Kaplan-Meier Plots", pp 21-31	- Cochran-Mantel-Haenszel chi-square test - log-rank test - Cox regression hazard ratio is weighted average of interval-specific risk ratios
	■ view recorded lecture "Ch 5-7 Kaplan-Meier probabilities" (14 min) ■ read Ch 5-7 "Intro Cox regression & Kaplan-Meier Plots", pp 31-38 ■ complete Lab_Nov20: Ch 5-97: ch 5-7 prob 4	- hazard rate - relative risk - Kaplan-Meier probabilities - cumulative survival probabilies - failure probabilities
	■ view recorded lecture "Ch 5-7 adjusted Kaplan-Meier graphs" (10 min)■ ■ read Ch 5-7 "Intro Cox regression & Kaplan-Meier Plots", pp 39-48 ■ complete Lab_ Nov20: Ch 5-97: ch 5-7 probs 5 to 9	- Kaplan-Meier survival graph - Kaplan-Meier failure graph - covariate-adjusted Kaplan-Meier graph
	■ view recorded lecture "Ch 5-7 truncated analysis" (6 min) ■ read Ch 5-7 "Intro Cox regression & Kaplan-Meier Plots", pp 48-51	- truncated survival analysis
Tue Nov 27	■ material covered today is not yet recorded, so you are not responsible for this material (it is still useful, so it is worth reading)	
Tue Dec 4	■ read Ch 2-15 "equivalence tests and noninferiority tests", pp 1-5 ■ no recorded lecture available	- equivalence test health disparity motivating example - published example of incorrectly concluding equivalence from nonsignificant p value - Hoenig and Heisey's published proof

		that lack of significance does not allow conclusion of equivalence or noninferiority - logical construct of a proof by contradiction
	■ read Ch 2-15 “equivalence tests and noninferiority tests”, pp 6-7 ■ no recorded lecture available	- post hoc power based on observed effect size was also refuted by Hoenig and Heisey’s published proof - example showing SPSS provides illogical post hoc power calculations
	■ read Ch 2-15 “equivalence tests and noninferiority tests”, pp 6-9 ■ no recorded lecture available	- FDA’s ±20% rule for equivalence testing of bioavailability of drugs - null and alternative hypotheses of equivalence testing - Schuirmann’s Two One-Sided Tests (TOST) Procedure
	■ read Ch 2-15 “equivalence tests and noninferiority tests”, pp 10-15 ■ no recorded lecture available	- confidence interval approach to equivalence testing - 90% CIs or 95% CIs - noninferiority tests using CIs - 2-sided vs 1-sided CIs controversy
	■ read Ch 2-15 “equivalence tests and noninferiority tests”, pp 16-17 ■ no recorded lecture available	- what researchers are using in published articles (only 23% do it correctly) - available software for equivalence testing
	■ read Ch 2-15 “equivalence tests and noninferiority tests”, pp 18-21 (still in draft stage—skip these pages) ■ no recorded lecture available	- example of noninferiority analysis based on proportions in Stata using CIs
	■ read Ch 2-15 “equivalence tests and noninferiority tests”, pp 22-23 ■ no recorded lecture available ■ In class discussion of Munger article (online students not responsible for this material)	- testing noninferiority and superiority in same study - Munger article example of testing noninferiorty and superiority in same study
	■ read Ch 2-15 “equivalence tests and noninferiority tests”, pp 24-32 (still in draft stage—skip these pages) ■ no recorded lecture available	- better CI formulas for difference between two proportions - *rdci* command
	■ read Ch 5-16 “correlated data: ANCOVA vs change scores”, pp 1-6 ■ no recorded lecture available	- paired sample t-test (two repeated measurements) - equivalence of paired sample t-test and one sample t-test on change scores

	■ read Ch 5-16 “correlated data: ANCOVA vs change scores”, pp 7-9 ■ no recorded lecture available	- designing a repeated measures experiment - crossover study design - parallel groups design - when is a paired t-test more powerful than an independent groups t-test
	■ read Ch 5-16 “correlated data: ANCOVA vs change scores”, pp 10-13 ■ no recorded lecture available	- three ways to analyze a parallel groups design study: post, change, ANCOVA - change approach more powerful than post approach only if r = 0.50
	■ read Ch 5-16 “correlated data: ANCOVA vs change scores”, pp 14-21 ■ no recorded lecture available	- example of three approaches
	■ read Ch 5-16 “correlated data: ANCOVA vs change scores”, pp 22-29 ■ no recorded lecture available	- how ANCOVA works - Monte Carlo simulation of three approaches - statistical methods suggestion for ANCOVA - repeated baseline measurements
	■ read Ch 5-29 “model risk difference and risk ratio”, pp 1-3 ■ no recorded lecture available	- odds ratio only approprimates a risk ratio when the rare disease assumption is met (when outcome occurs < 10%) - logistic regression only justified when rare disease assumption is met if odds ratio is to be reported (or when it is required in a case-control study) - prominent statisticians are pushing researchers to use alternative regression models that estimate risk ratio directly
	■ read Ch 5-29 “model risk difference and risk ratio”, pp 4-12 ■ no recorded lecture available	- log-binomial model - risk difference using *margins* command
	■ read Ch 5-29 “model risk difference and risk ratio”, pp 13-24 ■ no recorded lecture available	- binomial regression - Poisson regression - Poisson regresson can be used to model three different effect estimates: (1) counts of rare events, (2) rate, where rate = cases/(person time), and (3) risk, or incidence proportion, where risk = cases/N = cases/(number at risk).
	■ read Ch 5-29 “model risk difference and risk ratio”, pp 25-26 ■ no recorded lecture available	- responding to reviewers who want logistic regression

	■ read Ch 5-29 "model risk difference and risk ratio", pp 27-32 ■ no recorded lecture available	- using logistic regression and then obtaining risk ratio and risk difference using adjrr command
	■ midterm 3 assigned (covers Oct 24 to Nov 28 material) – end of finals week to complete	

Note: This syllabus is meant to serve as an outline and guide for our course. Please note that I may modify it with reasonable notice to you. I may also modify the Course Schedule to accommodate the needs of our class. Any changes will be announced in class and posted on Canvas under Announcements.