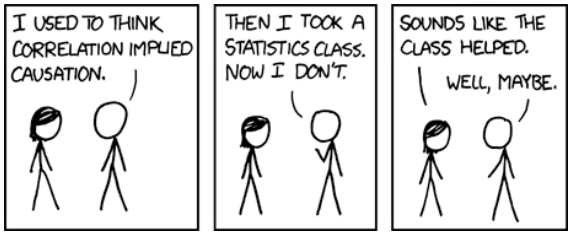
BIOS 6611
Biostatistical Methods I
Fall Semester 2018

Instructor: Alex Kaizer, Ph.D.	**Teaching Assistant:** Wenru Zhou
Office: Building 406, Room 109, (303) 724-4101	**Office:** Ed 1 1501
e-mail: alex.kaizer@ucdenver.edu	**e-mail:** wenru.zhou@ucdenver.edu
Office hours: Ed 1 1501, TBD, or by appointment	**Office hours:** W 2:30-4:00, Th 9:30-11:30, Th 4:30-6:00; or by appointment
 Source: xkcd.com	**Teaching Assistant:** Yaxu Zhuang
	Office: Ed 1 1501
	e-mail: yaxu.zhuang@ucdenver.edu
	Office hours: T 5:30-7:00, W 4:30-6:30; or by appointment
	Teaching Assistant: Liz Litkowski
	Office: Ed 1 1501
	e-mail: elizabeth.litkowski@ucdenver.edu
	Office hours: T 5:30-7:00, Th 4:30-6:00; or by appointment

COURSE GOALS

This is the first of a two-semester applied statistics sequence (3 credits per semester) designed primarily to equip graduate students with a practical knowledge of the quantitative methods most frequently used in medical research. This semester will emphasize an introduction to applied biostatistics, including probability models, simulation, estimation methods, sampling distributions, principles of hypothesis testing, modes of inference, resampling methods, linear regression models and the use of the R and SAS statistical packages to perform data simulation studies and data analysis. Concepts will be illustrated through the use of examples in the fields of medicine, biology, epidemiology and public health. Written and graphical presentation of methods and results, and interpretation of analytic results will be emphasized.

PREREQUISITES

Differential and integral calculus is a prerequisite for the course, along with an introductory course in applied probability and statistics. Previous coursework or experience with a statistical package (e.g. R, SAS, Stata, SPSS) is also assumed. The course is required for the MS in Biostatistics, the PhD and DrPH in Epidemiology, the PhD in Health Services Research and the DrPH in Environmental and Occupational Health. It is appropriate for other graduate students with strong mathematical preparation who may enroll with the stated prerequisites.

INTEGRATION OF THIS COURSE WITH OTHER BIOSTATISTICS COURSES

As an introductory course, this course provides a foundation for analytic methods covered in BIOS 6612, BIOS 6621, BIOS 6622, BIOS 6623, BIOS 6624, BIOS 6640, BIOS 6643, BIOS 6646, BIOS 6655 and BIOS 6660. Examples will be used from several areas including epidemiology, behavioral and community health, environmental health, health services, rehabilitation medicine, cancer biology and oncology.

COMPETENCIES MAPPED TO THIS COURSE FOR ASSESSMENT

This course partially or fully addresses the following MS core knowledge and competencies, and is used for assessing achievement.

Identifier	Competency	Description of Assessment
MS-BIOS 2	Apply statistical concepts of basic study designs including bias, confounding and efficiency, and identify strengths and weaknesses of experimental and observational designs.	Multiple written homeworks and exams
MS-BIOS 3	Carry out exploratory and descriptive analyses of complex data using standard statistical software and methods of data summary and visualization.	Multiple written homeworks and exams
MS-BIOS 4	Carry out valid and efficient modeling, estimation, model checking and inference using standard statistical methods and software.	Multiple written homeworks and exams
MS-BIOS 5	Demonstrate statistical programming proficiency, good coding style and use of reproducible research principles using leading statistical software.	Multiple written homeworks and project
PhD-EPID 10	Carry out appropriate power or precision calculations to ensure that sample size is sufficient to achieve the scientific aims or address a specific research hypothesis.	Multiple written homeworks and in-class examples

TEXTBOOKS AND RESOURCES FOR COURSE MATERIAL AND R/SAS SOFTWARE PROGRAMS

Optional course text:

For second half of course:
Kleinbaum, D.G., Kupper, L.L., Nizam, A., Rosenberg, E.S. Applied Regression Analysis and Other Multivariable Methods. Boston, MA: Brooks/Cole Cengage Learning, 2014. (4th edition, 2008 ok, too). https://www.cengagebrain.com/shop/isbn/9781285051086

For additional study/reference:

Chihara, L. and Hesterberg, T. Mathematical Statistics with Resampling and R. Hoboken, NJ: Wiley and Sons, 2011. Some used copy prices ok at Amazon and Barnes and Noble.

Diez, D.M., Barr, C.D., Mine Çetinkaya-Rundel. Introductory Statistics with Randomization and Simulation, OpenIntro CreateSpace Independent Publishing Platform.
Free download at https://www.openintro.org/stat/textbook.php?stat_book=os

Rosner, B. Fundamentals of Biostatistics, eighth (or seventh or sixth edition). Pacific Grove, CA: Duxbury, 2015 (2011, 2005).

Vittinghoff, E., Glidden, D.V., Shiboski, S.C. and McCulloch, C. Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models, 1st Edition or 2nd Edition. New York: Springer, 2005, 2012. eBook rentals at http://rentals.springer.com/product/9781461413530

Useful Websites and Resources for learning and using R:

Quick R http://www.statmethods.net/
- Provides examples for basic syntax, manipulating data, producing charts, and more sophisticated analyses as well as links out to more complete resources.

Cookbook for R http://www.cookbook-r.com/
- Based off a book in the O'Reilly series (a great book series on all programming topics), this site provides templates and code examples, as well as great templates for ggplot2, an advanced graphing library for R.

R Studio http://www.rstudio.com/
- R Studio provides a full-featured "IDE" (Integrated Design Environment) for R that features project management, plot history, exporting and importing of data sets and plots.

UCLA – Statistics – Resources for learning R https://stats.idre.ucla.edu/r/
- Lots of great data examples, tutorials, and interpretations of methods in R

Useful Websites for SAS:

SAS University Edition Download (free installation of SAS)
https://www.sas.com/en_us/software/university-edition/download-software.html

UCLA – Statistics – Resources for learning SAS https://stats.idre.ucla.edu/sas/
- Lots of great data examples, tutorials, and interpretations of methods in SAS

More learning resources:
SASCRUNCH.COM - http://www.sascrunch.com/articles.html
StatTag for Reproducible Research using SAS - http://sites.northwestern.edu/stattag/

Books on Software:

Cody, R. Learning SAS by Example: A Programmer's Guide. Cary, NC: SAS Institute, 2007.

Elliott, R.J and Morrell, C.H. Learning SAS in the Computer Lab, third edition. Boston, MA: Brooks/Cole Cengage Learning, 2010.

Kleinman, K. and Horton, N.J. SAS and R: Data Management, Statistical Analysis, and Graphics, second edition. Boca Raton, FL: Chapman and Hall/CRC, 2014. These authors also wrote separate R and SAS books.

Delwiche, L.D. and Slaughter, S.J. The Little SAS Book – a primer, fifth edition. Cary, NC: SAS Institute, 2012.

O'Rourke, N., Hatcher, L. and Stepanski, E.J. A Step-by-step Approach to Using SAS for Univariate and Multivariate Statistics, second edition. Cary, NC: SAS Institute, 2012.

BIOS 6611: Biostatistical Methods I
STUDENT EVALUATION

The course will follow a reading, lecture, problem-solving and discussion format. Enthusiastic class participation is expected! This should be in the form of asking and answering questions. No question is too basic and all answers will be respectfully received by the instructor, TAs and fellow students. Each class session will be recorded and posted to Canvas for you to review your notes, if you wish, or if you are not able to attend class. You will be responsible for the material covered in the in-person class sessions.

Your evaluation in the course will be based in part on several homework assignments that will include both application problems using statistical software and mathematical exercises. You are encouraged to work together on these assignments, *but the work you hand in must be your own*. Grading will be done on a select subset of each homework assignment, but you must submit a complete assignment to be graded. There will be two in-class exams, a midterm (1.5 hours long) and a final (3 hours long). Finally, you will be required to complete a simulation plus data analysis project by the end of the semester; more information will be provided about halfway through the course. The weight to each course component is:

Component	**% of Grade**
Homework	35%
Midterm Exam	22.5%
Final Exam	22.5%
Simulation + Data Analysis Project	15%
Class Participation	5%

You are expected to read the course materials, log into the Canvas and use the materials provided there, and use the interactive learning tools identified in the syllabus. All assignments and the project should be organized neatly and as much material as possible should be typed. Homework assignments and the project must be completed and turned in ***on time*** as required. Failure to accomplish this will result in a grade of Incomplete. These should be submitted electronically through Canvas.

In-class exams are scheduled at approximately mid-semester and at the end of the semester. ***These are on the honor system (see p. 13) and are to reflect only your own work without the aid of others.*** The first exam will cover material presented up to that point in the course. The final exam will not directly cover the material from the first half of the semester but will depend on a cumulative understanding of all course material.

A final grade will be assigned on a fixed (not "curved" scale):

A	94.00-100	B-	80.00-82.99	D+	67.00-69.99
A-	90.00-93.99	C+	77.00-79.99	D	63.00-66.99
B+	87.00-89.99	C	73.00-76.99	D-	60.00-62.99
B	83.00-86.99	C-	70.00-72.99	F	0 – 59.99

BIOS 6611: Biostatistical Methods I – Fall 2018
TTh 2:30-3:50 PM - L28 (Ed 2 South): Rm 1307
Tentative Schedule at a Glance

Week	Date	Lecture	Topic	Readings	Homework Due (by Noon)
1	8/28	1	Intro/Overview/Sampling/Simulation	Notes	9/4 (L1-2)
	8/30	2	Central Limit Theorem	Notes, Rosner 6, C&H 4	
2	9/4	3	Expected value, variance Bias, efficiency, consistency	Notes, Rosner 4-5, C&H 6	9/17 (L3-4)
	9/6	4	Discrete Distributions: Binomial, Poisson	Notes, Rosner 4	
3	9/11	5	Continuous Distributions: Normal	Notes, Rosner 5	9/24 (L5-6)
	9/13	6	Functions of random variables	Notes, Rosner 5	
4	9/18	7	Hypothesis testing – Fisher vs. N-P	Notes, Papers	10/1 (L7-8)
	9/20	8	Power and sample size	Notes, Rosner 7	
5	9/25	9	Conditional probability	Notes, Rosner 3	10/8 (L9-10)
	9/27	10	2x2 Tables: Effect measures, Tests	Notes, Rosner 3, 13	
6	10/2	11	Bayesian inference	C&H 10, Rosner 7, GUSTO papers	10/15 (L11-12)
	10/4	12	Frequentist inf: Permutation & rank tests	C&H 3, Rosner 9	
7	10/9	13	Bootstrap sampling	C&H 5, Paper	10/22 (L13-15)
	10/11	14	ANOVA and GLM	Notes, Rosner 12	
8	10/16	15	Multiple comparisons/tests	Notes, Rosner 12	
	10/18		Mid-term Exam		
9	10/23	16	Simple Linear Regression (SLR)	KKNR 1-4,5	11/5 (L16-18)
	10/25	17	SLR: Statistical Inference		
10	10/30	18	SLR: Example	KKNR 7	
	11/1	19	Multiple Linear Regression (MLR)		
11	11/6	20	Matrix Approach to LR	KKNR Appx B	Bonus Hwk 11/12
	11/8	20	Matrix Approach to LR	8-9	
12	11/13	21	MLR: Confounding and Mediation	KKNR 11	11/26 (L19,21-22)
	11/15	22	MLR: MLE and Effect Modification	KKNR 11-12, 21	
13	11/20	23	MLR: Categorical Predictors	KKNR 12-13	12/3 (L23-24)
	11/22		Thanksgiving – no class		
14	11/27	24	MLR: Testing General Linear Hypotheses	Notes	Practice Hwk (L25-26)
	11/29	25	MLR: Polynomial Regression	KKNR 15	
15	12/4	26	MLR: Regression Diagnostics	KKNR 14	
	12/6	Review	Review		
	12/7		Data Analysis Project Due		
16	12/10-14		Exam Week		
	12/11		Final Exam (1:00-4:00 PM)		
	12/19		Grades Due		

TENTATIVE **Lecture Topics**

DATE	**TOPIC**
Week 1	Introduction and overview – inference from data
8/28	Framing based on Algorithms vs. inference: Efron and Hastie, *Computer Age Statistical Inference: Algorithms, Evidence, and Data Science*, IMS Monographs, 2016 Models vs. algorithms: Breiman, *Statistical Modeling: The Two Cultures*, Stat Sci, 2001 Generating random numbers, using them for randomized experiments Pseudo-random numbers, tables, functions in R Random sampling from theoretical distributions Simulation for understanding – properties of estimators Simulation for experimentation – testing hypotheses http://www4.stat.ncsu.edu/~davidian/st810a/simulation_handout.pdf Goodman (1999) paper on p-values Benjamin et al. (2017) paper on p-values
8/30	Frequentist Inference: From Random Samples to Point estimates to Sampling distributions The Mean and its Standard Error The Central Limit Theorem Sampling distribution of the sample variance
Week 2	
9/4	Expected value and Bias: Mean vs. median vs. mode, sample variance Properties of estimators: Consistency, (relative) efficiency, mean square error Types of estimators: method of moments, maximum likelihood, ordinary least squares
9/6	Common probability models Discrete distributions: Binomial and Poisson
Week 3	
9/11	Continuous distributions: Normal, and approximations
9/13	Functions of random variables Expected values and variances of sums and products of independent random variables

Week 4 Frequentist inference – hypothesis testing

9/18
R.A. Fisher and significance testing
Neyman-Pearson and hypothesis testing
Type I and II errors, likelihood ratios, significance and confidence, p-values
Goodman (1999) paper on Bayes factors

9/20 Frequentist inference: sample size and power

Week 5

9/25
Decision-making with conditional probabilities
Sensitivity, Specificity, ROC Curves, likelihood ratios
Predictive value and Bayes' Rule
Goodman (1999) paper on Bayes factors

9/27 2x2 Tables – Study designs, effect measures, hypothesis tests

Week 6

10/2
Bayesian inference
NEJM (Frequentist) and JAMA (Bayes) papers on the GUSTO clinical trial

10/4
Frequentist nonparametric inference
Nonparametric hypothesis tests:
- Permutation tests: Sampling without replacement
- Rank-based methods: Sign, Wilcoxon signed rank, Wilcoxon/Mann-Whitney tests

Week 7

10/9
More frequentist inference based on resampling
- The nonparametric bootstrap - Sampling with replacement

Tim Hesterberg paper on n>=30 rule of thumb
Tim Hesterberg review paper on bootstrap sampling

10/11
Frequentist parametric inference: Comparing more than two means
ANOVA and Introduction to the general linear model

Week 8

10/16
Multiple comparisons, linear contrasts
Multiple testing and the false discovery rate
- Example: Reproducibility and regression to the mean example from Westfall/Ioannidis

10/18 Midterm Exam

Week 9	Frequentist parametric inference: Linear regression models
10/23	Simple Linear Regression (SLR)
10/25	SLR: Statistical Inference
Week 10	
10/30	SLR: Example
11/1	Multiple Linear Regression (MLR)
Week 11	
11/6	Matrix Approach to Linear Regression
11/8	Matrix Approach to Linear Regression
Week 12	
11/13	MLR: Confounding and Mediation
11/15	MLR: MLE and Effect Modification
Week 13	
11/20	MLR: Categorical Predictors
11/22	Thanksgiving – no class
Week 14	
11/27	MLR: Testing General Linear Hypotheses
11/29	MLR: Polynomial Regression
Week 15	
12/4	MLR: Regression Diagnostics
12/6	Review
12/7	**Final Project Due**
Week 16 (Finals Week)	
12/11	Final exam **1:00-4:00 PM**

BIOS 6611
Biostatistical Methods I
Fall Semester 2018
Class and Room Schedule

LECTURES:

Tuesdays	2:30-3:50 PM	Ed 2 South L28-Rm 1307
Thursdays	2:30-3:50 PM	Ed 2 South L28-Rm 1307

Course handouts, assignments and datasets will be available on the course website on Canvas: https://ucdenver.instructure.com/login. For further information on Canvas, be sure to consult the Canvas Student Guide: http://guides.instructure.com/m/4212.

Be sure to check the website and your e-mail account regularly for updates, changes and course communication! ***You will be responsible for bringing any course materials to class (either hardcopy or e-copy); no hardcopies will be provided, unless notified otherwise by the instructor.***

COMPUTER LAB:

All office hours will be held in **P26-1501 in the Ed 1 building**. There is also a computer lab located in P28-2201C in the Ed 2 North building. These rooms are accessible 24 hours a day, 7 days a week, using your student ID badge when not reserved for specific classes. To check availability of the rooms at times other than scheduled office hours, you can browse the schedule at: http://schedule.ucdenver.edu/virtualems/browseforspace.aspx

STATISTICAL SOFTWARE DOWNLOADS

R and RStudio
You can download these at: https://cran.r-project.org/ and https://www.rstudio.com/.

SAS
For SAS University, a free installation of SAS, the download link is:
https://www.sas.com/en_us/software/university-edition/download-software.html

PURCHASE OF SAS SOFTWARE LICENSE (optional):
SAS licenses for 2018-2019 are now available, for $90 per user. If you can purchase using a credit card or Purchase card, make your purchase through the CU Book Store website (can take up to two days to receive license information for SAS):
https://www.cubookstore.com/p-90170-sas-statistical-software.aspx

UCD/CSPH ACADEMIC HONOR CODE

Academic Conduct Policy

All students are expected to familiarize themselves with and to abide by the Honor Code of the Colorado School of Public Health. The Student Honor Code that can be found at http://www.ucdenver.edu/academics/colleges/PublicHealth/resourcesfor/currentstudents/academics/Documents/PoliciesHandbooks/CSPH_Honor_Code.pdf or the Student Resources Section of the CSPH website.

All aspects of this course are conducted under the honor system. All students should have developed the qualities of honesty and integrity, and each student should apply these principles to his or her academic and subsequent professional career. All students are expected to have achieved a level of maturity, which is reflected in appropriate conduct at all times. Related to academic honesty, all work done on exams and other assignments is to be done independently, unless specific instruction to the contrary is provided by the course instructor.

Any student found to have committed acts of misconduct including, but not limited to cheating, plagiarism, misconduct of research, breach of confidentiality, or illegal or unlawful acts) will be subject to the procedures outlined in the CSPH Honor Code.

You are required to take the CU Online course on Academic Integrity and submit your completion certificate. You'll find information under the module called Course Home -> Academic Integrity and the Assignment basket -> Academic Integrity Certification.

For every exam in this course you will be asked to sign the following statement:

I understand that my participation in this examination and in all academic and professional activities as a UCD student is bound by the provisions of the UCD Honor Code. I understand that work on this exam and other assignments is to be done independently unless specific instruction to the contrary is provided.

On my honor, I have neither given nor received aid on this examination.

_____ _____

Signature Date

Disability Policy

Students requesting accommodation should contact the Office of Disability Resources and Services. Their staff will assist in determining reasonable accommodations as well as coordinating the approved accommodations.

Sherry Holden | Coordinator
University of Colorado Anschutz Medical Campus Disability Resources & Services
Bldg. 500, Room Q20-EG 305A
Phone: (303) 724-5640, Fax (303) 724-5641
Part-time: Monday, Tuesday and Thursday
sherry.holden@ucdenver.edu

Selim Özi | Assistive Technology Specialist, Accommodation Coordinator
University of Colorado Anschutz Medical Campus Disability Resources & Services
Mail Stop A010, Building 500, Room Q20-EG 306
Phone: (303) 724 8428, Fax: (303) 724 5641
selim.ozi@ucdenver.edu

Be aware that the determination of accommodations can take a long period of time. No accommodations will be made for the course until written documentation is provided by the Disability resources and services office to the course directors. It is the student's responsibility to coordinate approved accommodations with the Disability resources and services office in advance.

Further general Information regarding disability resources and services can be found at:
http://www.ucdenver.edu/student-services/resources/disability-resources-services/accommodations/Pages/accommodations.aspx

Students can set up an appointment at:
http://www.ucdenver.edu/student-services/resources/disability-resources-services/about-office/contact-us-CUAnschutz/Pages/form.aspx

Other Resources

AMC Student Mental Health Service
Comprehensive and confidential mental health services available for all students enrolled in the schools located on the Anschutz Medical Campus (Medical, Dental, Pharmacy, Public Health, Nursing, Physician Assistant, Physical Therapy, Graduate School, etc.). Initial appointments are scheduled relatively quickly, generally within the same week. Can call 303-724-4716, email SMHservice@ucdenver.edu, or visit their website for more details:
http://www.ucdenver.edu/academics/colleges/medicalschool/departments/psychiatry/PatientCare/StudentMentalHealth/Pages/Student-Mental-Health-Service.aspx

University of Colorado Anschutz Medical Campus Phoenix Center (PCA)
PCA serves the entire campus community, including students, faculty, and staff. They provide free, confidential support services and resource referrals for issues of interpersonal violence (an umbrella term that encompasses relationship violence, sexual violence, and stalking). Located in **Education 2 North room 5232** they have office hours Monday-Friday from 8 am-5 pm. For appointments or information call 303-724-9120. They also have a **24/7 Free and Confidential Helpline at 303-556-CALL (2255)** which can provide information on interpersonal violence, referrals, options, and next steps.
http://www.ucdenver.edu/anschutz/studentresources/student-assistance/student-resources/Pages/CU-Anschutz-Advocacy-and-Support-Center.aspx

The Writing Center at CU – AMC:
If you would like assistance in improving your writing or homework and/or mini project assignments, The Writing Center is a great on-campus resource. Be sure to look into their services:
http://www.ucdenver.edu/academics/colleges/CLAS/Centers/writing/Pages/TheWritingCenter.aspx

SpringerLink
Students have access to download PDF or EPUB versions of many Springer published textbooks and articles for free through SpringerLink (select CU Denver campus and log into Passport):
https://library.auraria.edu/databases/springerlink

_____ ∞ _____

ACKNOWLEDGMENTS

Thank you to the following individuals for their valuable contributions to the syllabus, lectures, lab handouts, assignments, and other course materials:

Anna Barón
Mike Daniels
Patrick Hosokawa
Cuining (Choo) Liu
Mikaela Miller
Molly Nowels
Sarah Ryan
Yu Zhang
Melanie Bell
Bob Engle
John Kittelson
Sharon Lutz
Camille Moore
Lorri Ogden
Stephanie Santorico
Eleanor Cotton
Gary Grunwald
Judy Koslov
Susan Mikulich
Jamie Nelson
Katie Roberts
Pam Wolfe