SYLLABUS & COURSE INFORMATION

SCHOOL OF PUBLIC HEALTH
UNIVERSITY OF MINNESOTA

PUBH 7445

Statistics for Human Genetics and Molecular Biology

Fall 2019

COURSE & CONTACT INFORMATION

Credits: 3
Meeting Day(s): MWF
Meeting Time: 1:25-2:15
Meeting Place: Mayo C381

Instructor: Cavan Reilly
Email: cavanr@biostat.umn.edu
Office Phone: 612-624-9644
Fax: 612-626-0660
Office Hours: MW 2:30-3:30
Office Location: 359 University Office Plaza

Teaching Assistant: Souvik Seal
Email: sealx017@umn.edu
Office Hour: W 1-2
Office Location: A446 Mayo

COURSE DESCRIPTION

This master's-level course focuses on quantitative skills and knowledge for interdisciplinary applications in genetics and molecular biology. Topics include gene association studies, analyses of gene expression data, analysis of next generation sequence data, false discovery rate estimation, gene set enrichment analysis and machine learning algorithms. Students will conduct extensive data analysis using R/Bioconductor.

COURSE PREREQUISITES

Pub 6450-6451 or equivalent or permission of instructor. Some background with molecular biology is desirable but not required.

COURSE GOALS & OBJECTIVES

By the end of the course, students should have a basic knowledge of statistical tools applied to human genetics, basic R/Bioconductor programming skills and substantial experience applying these skills and experience interpreting results and communicating these results to others.

METHODS OF INSTRUCTION AND WORK EXPECTATIONS

Course Workload Expectations

PUBH 7445 is a 3-credit course. The University expects that for each credit, you will spend a minimum of three hours per week attending class or comparable online activity, reading, studying, completing assignments, etc. over the course of a 15-week term. Thus, this course requires approximately 135 hours of effort spread over the course of the term in order to earn an average grade.

COURSE TEXT & READINGS

John Verzani's SimpleR notes (https://cran.r-project.org/doc/contrib/Verzani-SimpleR.pdf)
Foulkes (2009): Applied Statistical Genetics with R
Hahne, Huber, Gentleman, and Falcon (2008): Bioconductor Case Studies

COURSE OUTLINE/WEEKLY SCHEDULE

Week 1: Introduction/review of basic concepts of genetics and molecular biology
Week 2: Introduction/review of basic statistics and the use of R
Week 3: Linkage disequilibrium and Hardy Weinberg equilibrium
Week 4: Multiple comparisons
Week 5: Haplotype estimation
Week 6: Classification/supervised learning with categorical predictors
Week 7: More on classification/supervised learning
Week 8: Introduction to microarray data analysis
Week 9: Inference for differential expression
Week 10: More on inference for differential expression
Week 11: Next generation sequencing: bioinformatics
Week 12: Next generation sequencing: biostatistics
Week 13: Machine learning approaches to gene expression data
Week 14: R tools for pathway analysis in RNA-Seq data
Week 15: Proteomics and metabolomics

SPH AND UNIVERSITY POLICIES & RESOURCES

The School of Public Health maintains up-to-date information about resources available to students, as well as formal course policies, on our website at www.sph.umn.edu/student-policies/. Students are expected to read and understand all policy information available at this link and are encouraged to make use of the resources available.

The University of Minnesota has official policies, including but not limited to the following:

- Grade definitions
- Scholastic dishonesty
- Makeup work for legitimate absences
- Student conduct code
- Sexual harassment, sexual assault, stalking and relationship violence
- Equity, diversity, equal employment opportunity, and affirmative action
- Disability services
- Academic freedom and responsibility

Resources available for students include:

- Confidential mental health services
- Disability accommodations
- Housing and financial instability resources
- Technology help
- Academic support

EVALUATION & GRADING

Performance on homework assignments (70%)) and the final course project (30%).

Late homework will lose 30% of the total points each day it is late, unless prior arrangements have been made. Homework will not be accepted after the time at which graded homework are returned.

Grading Scale
The University uses plus and minus grading on a 4.000 cumulative grade point scale in accordance with the following, and you can expect the grade lines to be drawn as follows:

% In Class	Grade	GPA
93 - 100%	A	4.000
90 - 92%	A-	3.667
87 - 89%	B+	3.333
83 - 86%	B	3.000
80 - 82%	B-	2.667
77 - 79%	C+	2.333
73 - 76%	C	2.000
70 - 72%	C-	1.667
67 - 69%	D+	1.333
63 - 66%	D	1.000
< 62%	F	

- A = achievement that is outstanding relative to the level necessary to meet course requirements.
- B = achievement that is significantly above the level necessary to meet course requirements.
- C = achievement that meets the course requirements in every respect.
- D = achievement that is worthy of credit even though it fails to meet fully the course requirements.
- F = failure because work was either (1) completed but at a level of achievement that is not worthy of credit or (2) was not completed and there was no agreement between the instructor and the student that the student would be awarded an I (Incomplete).
- S = achievement that is satisfactory, which is equivalent to a C- or better
- N = achievement that is not satisfactory and signifies that the work was either 1) completed but at a level that is not worthy of credit, or 2) not completed and there was no agreement between the instructor and student that the student would receive an I (Incomplete).

Evaluation/Grading Policy	Evaluation/Grading Policy Description
Scholastic Dishonesty, Plagiarism, Cheating, etc.	You are expected to do your own academic work and cite sources as necessary. Failing to do so is scholastic dishonesty. Scholastic dishonesty means plagiarizing; cheating on assignments or examinations; engaging in unauthorized collaboration on academic work; taking, acquiring, or using test materials without faculty permission; submitting false or incomplete records of academic achievement; acting alone or in cooperation with another to falsify records or to obtain dishonestly grades, honors, awards, or professional endorsement; altering, forging, or misusing a University academic record; or fabricating or falsifying data, research procedures, or data analysis (As defined in the Student Conduct Code). For additional information, please see https://z.umn.edu/dishonesty The Office for Student Conduct and Academic Integrity has compiled a useful list of Frequently Asked Questions pertaining to scholastic dishonesty: https://z.umn.edu/integrity. If you have additional questions, please clarify with your instructor. Your instructor can respond to your specific questions regarding what would constitute scholastic dishonesty in the context of a particular class-e.g., whether collaboration on assignments is permitted, requirements and methods for citing sources, if electronic aids are permitted or prohibited during an exam. Indiana University offers a clear description of plagiarism and an online quiz to check your understanding (http://z.umn.edu/iuplagiarism).

CEPH COMPETENCIES

Competency	Learning Objectives	Assessment Strategies
Select quantitative and qualitative data collection methods appropriate for a given health context	Apply appropriate statistical estimation techniques to answer scientific questions using correlated data. Understand both the strengths and limitations of these techniques.	Homework assignments Final project
Analyze quantitative and qualitative data using biostatics, informatics, computer-based programming and software as appropriate	Describe and summarize datasets both graphically and numerically. Apply appropriate statistical estimation techniques to answer scientific questions.	Homework assignments Final project
Communicate audience-appropriate public health content, both in writing and through oral presentation	Write solutions to homework exercises.	Homework assignments Final project