SYLLABUS & COURSE INFORMATION

SCHOOL OF PUBLIC HEALTH
UNIVERSITY OF MINNESOTA

PUBH 8401, SECTION 001

Linear Models

Fall 2022

COURSE & CONTACT INFORMATION

Credits: 3
Meeting Day(s): **MW**
Meeting Time: **9:45-11:00**
Meeting Place: **Nils Hasselmo Hall 2-101**

Instructor: **Drs. Xiaoou Li & Cavan Reilly**
Email: **lixx1766@umn.edu, cavanr@biostat.umn.edu**
Office Phone: **612-624-9644 (Cavan)**
Fax:
Office Hours: **TBD**
Office Location: **CCBR 359 (Cavan), Ford 347 (Xiaoou)**

COURSE DESCRIPTION

This course is concerned with the theory and application of linear models. The first part of the course will focus on general linear model theory from a coordinate-free geometric approach. The second half of the course covers theory, applications and computing for linear models, and concentrates on modeling, computation and data analysis. It is intended as a core course for biostatistics PhD students and statistics PhD students.

Acknowledgments

The contents of PubH 8401 have been developed with the contributions of numerous instructors. Drs. Xiaoou Li and Cavan Reilly, the current instructor, has been involved with the majority of recent content and modifications. Former faculty/instructors, including Drs. Yuhong Yang, Baolin Wu, Weihua Guan and Galin Jones, all had roles in either the conceptual development or actual content of the current course, and are acknowledged for their contributions.

COURSE PREREQUISITES

PubH 7405 and (concurrent registration of) Stat 8101 are required; or permission by instructors. It is assumed students have had calculus and are familiar with linear algebra. Please see the instructor if you have questions about the suitability of your background.

COURSE GOALS & OBJECTIVES

At the completion of this course, the student will understand the theoretical foundations for general linear models widely used in statistics, and be able to apply the appropriate modeling and analysis techniques for solving practical problems in health science related fields.

METHODS OF INSTRUCTION AND WORK EXPECTATIONS

Course Workload Expectations

PUBH 8401 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.

Learning Community

School of Public Health courses ask students to discuss frameworks, theory, policy, and more, often in the context of past and current events and policy debates. Many of our courses also ask students to work in teams or discussion groups. We do not come to our courses with identical backgrounds and experiences and building on what we already know about collaborating, listening, and engaging is critical to successful professional, academic, and scientific engagement with topics.

In this course, students are expected to engage with each other in respectful and thoughtful ways.

In group work, this can mean:

- Setting expectations with your groups about communication and response time during the first week of the semester (or as soon as groups are assigned) and contacting the TA or instructor if scheduling problems cannot be overcome.
- Setting clear deadlines and holding yourself and each other accountable.
- Determining the roles group members need to fulfill to successfully complete the project on time.
- Developing a rapport prior to beginning the project (what prior experience are you bringing to the project, what are your strengths as they apply to the project, what do you like to work on?)

In group discussion, this can mean:

- Respecting the identities and experiences of your classmates.
- Avoid broad statements and generalizations. Group discussions are another form of academic communication and responses to instructor questions in a group discussion are evaluated. Apply the same rigor to crafting discussion posts as you would for a paper.
- Consider your tone and language, especially when communicating in text format, as the lack of other cues can lead to misinterpretation.

Like other work in the course, all student to student communication is covered by the Student Conduct Code (https://z.umn.edu/studentconduct).

COURSE TEXT & READINGS

Charles McCulloch, Shayle Searle and John Neuhaus (2008). Generalized, Linear, and Mixed Models. 2nd Ed, Wiley. (E-Book link: https://ebookcentral.proquest.com/lib/umn/detail.action?docID=819128#goto_toc)

Michael J. Wichura (2009) The Coordinate-Free Approach to Linear Models, Cambridge University Press.

Martin J. Wainwright (2019) High-Dimensional Statistics: A Non-Asymptotic Viewpoint. Cambridge University Press.

Notes by R. D. Cook, K. Larntz, and S. Weisberg

Week	Topic	Readings	Activities/Assignments
Week 1 (Sep. 5)	• **Review of Linear Algebra: Linear Subspaces, Linear Transformations, Projections, Inner Products, Orthogonality, Coordinates with respect to an Orthonormal Basis, Orthogonal Projections, Orthogonal transformations, Matrices, Eigenvectors and Eigenvalues, Matrix Decompositions**	• **Notes by Cook, Larntz, and Weisberg**	•
Week 2 (Sep. 12)	• **Review of Linear Algebra (continued)**	• **Notes by Cook, Larntz, and Weisberg**	• **Homework due**
Week 3 (Sep. 19)	• **Linear Models: Estimation, Best Estimators, Gauss-Markov theorem, Estimability, Linear Restrictions, Generalized Least Squares, OLS vs Generalized Least Squares**	• **Notes by Cook, Larntz, and Weisberg** • **McCulloch, Searle, and Neuhaus, Chap. 1-4**	• **Homework due**
Week 4 (Sep. 26)	• **Linear Models (continued)**	• **Notes by Cook, Larntz, and Weisberg**	• **Homework due**
Week 5 (Oct. 3)	• **Linear Models (continued)**	• **Notes by Cook, Larntz, and Weisberg** • **McCulloch, Searle, and Neuhaus, Chap. 1-4**	• **Homework due**
Week 6 (Oct. 10)	• **Distribution Theory: Consistency of Least Squares, Characteristic Functions, Multivariate Normal, Chi-Squared and Quadratic forms, and F distributions**	• **Notes by Cook, Larntz, and Weisberg** • **McCulloch, Searle, and Neuhaus, Chap. 1-4**	• **Homework due**
Week 7 (Oct. 17)	• **Distribution Theory (continued)**	• **Notes by Cook, Larntz, and Weisberg** • **McCulloch, Searle, and Neuhaus, Chap. 1-4**	• **Homework due**
Week 8 (Oct. 24)	• **Inference:** • **Log-likelihood, Hypothesis testing, Geometry of F tests, Likelihood ratio tests, General Coordinate Free hypotheses, Parametric hypotheses, Relation of least squares estimators under NH and AH, Analysis of Variance Tables, F tests and t tests, Power and Sample Size, Simultaneous confidence intervals (Scheffe method, Bonferroni method, Tukey's method,**	• **Notes by Cook, Larntz, and Weisberg** • **McCulloch, Searle, and**	• **Midterm exam**

	and related methods), Multiple testing, Score tests for linear model diagnostics, admissibility and minimax estimation.	**Neuhaus, Chap. 1-4**	
Week 9 (Oct. 31)	• **Inference (continued)**	• **Notes by Cook, Larntz, and Weisberg** • **McCulloch, Searle, and Neuhaus, Chap. 1-4**	•
Week 10 (Nov. 7)	• **Inference (continued)**	• **Notes by Cook, Larntz, and Weisberg** • **McCulloch, Searle, and Neuhaus, Chap. 1-4**	• **Homework due**
Week 11 (Nov. 14)	• **Generalized Linear Models:** **Properties of exponential families; Estimation: IWLS, MLE; Consistency and efficiency; Hypothesis test: LRT, Wald, score test; Goodness-of-fit: deviance, Pearson X2; Overdispersion**	• **McCulloch, Searle, and Neuhaus, Chap. 5**	• **Homework due**
Week 12 (Nov. 21)	• **Generalized Linear Models (continued)**	• **McCulloch, Searle, and Neuhaus, Chap. 5**	• **Homework due**
Week 13 (Nov. 28)	• **Model Selection Theory: Bias-Variance tradeoff; AIC, BIC and related criteria, and associated properties; Ridge regression and Lasso; sparse linear approximations for high-dimensional linear models and minimax theory**	• **Notes by Reilly**	• **Homework due**
Week 14 (Dec. 5)	• **Model Selection Theory (continued)**	• **Notes by Reilly**	• **Homework due**
Week 15 (Dec. 12)	• **Model Selection Theory (continued)**	• **Notes by Reilly**	• **Homework due**

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

The course grade is based on midterm exam (30%), homework (30%), and a final exam (40%).

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).
Late Assignments	Late assignments are only allowed if a request is made prior to the due date. Details of how such requests are handled will be determined on a case by case basis.
Attendance Requirements	Attendance is not required.
Extra Credit	There is no opportunity for extra credit.

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. Understand both the strengths and limitations of these techniques.	Homework assignments Midterm and final exam
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 Midterm and final exam
Communicate audience-appropriate public health content, both in writing and through oral presentation	Write solutions to homework exercises.	Homework assignments Midterm and final exam