PUBH 8492, SECTION 001

Theories of Hierarchical and Other Richly Parameterized Linear Models
Fall 2023

COURSE & CONTACT INFORMATION

Credits: 3
Meeting Day(s): M/W
Meeting Time: 1:00 PM to 2:15 PM
Meeting Place: University Office Plaza, room 2-240 (The Ballroom)

Instructor: James S. Hodges, Professor, Division of Biostatistics
Email: hodge003@umn.edu
Office Phone: 612-626-9626
Office Hours: By appointment; e-mail is the best way to contact me.
Office Location: University Office Plaza Room 2-223 (at the west end of the 2nd floor)

COURSE DESCRIPTION

Linear richly-parameterized models include hierarchical models, dynamic linear models, linear mixed models, random regressions, smoothers (spatial and otherwise), longitudinal models, time series models, and many others. Existing theories are mainly schemes for specifying and fitting many such models. This course's first part describes the mixed-linear-model scheme (or syntax), reviews standard theory and computing for that scheme, and briefly reviews two other schemes. Apart from this, richly-parameterized models have nothing like the deep, powerful theory of ordinary linear models. This course's second part begins with the little theory that exists, then explores odd, surprising, or undesirable results that arise routinely in applying these models to real datasets. The purpose is to illustrate the problems a theory must address, emphasizing open questions and possible approaches to solving them. Class sessions are lectures by the instructor except that each student will present one lecture that they choose. The grade depends mostly on a final class project, which can be on any topic the instructor and student agree on, preferably related to their dissertation or research-assistant work. Students will present their projects in class and hand in a written version.

COURSE PREREQUISITES

Stat 8101/8102 (or its equivalent), and Stat 8311 or PubH 8400 (Linear Models), and Biostatistics or Statistics PhD student, or permission of instructor.

COURSE GOALS & OBJECTIVES

Present a theoretical approach to the large class of richly-parameterized linear models. Survey the beginnings of a theory of these models using datasets mostly from the instructor's collaborative projects to describe unsolved and partly-solved problems and approaches to working those problems.

METHODS OF INSTRUCTION AND WORK EXPECTATIONS

Course Workload Expectations

Class sessions consist of lectures by the instructor and one lecture by each student, and presentations of students' class projects. Although there are some homework exercises, students' grades depend mostly on the final class project presented orally (in class) and in writing. Each student selects the project's topic, subject to the instructor's approval and preferably related to their dissertation or research-assistant work. Each student will meet two or three times with the instructor about their project, depending on the number of students enrolled. Students are expected to read extensively from the course texts and from assigned journal articles. All articles for this course are on the instructor's web page (http://www.biostat.umn.edu/~hodges/, third item under "Links").

All student-to-student communication is covered by the Student Conduct Code.

COURSE TEXT & READINGS

This course uses journal articles. It is good practice to use a citation manager to keep track of your readings. More information about citation managers is available at https://www.lib.umn.edu/pim/citation.

Required text: Ruppert D, Wand MP, Carroll RJ (2003). *Semiparametric Regression*. Cambridge U Press

Supplementary texts: Hodges JS (2014). *Richly Parameterized Linear Models: Additive, Time Series, and Spatial Models Using Random Effects*, Chapman and Hall; Rue H, Held L (2005). *Gaussian Markov Random Fields: Theory and Applications*, Chapman and Hall; Lee Y, Nelder JA, Pawitan Y (2006). *Generalized Linear Models with Random Effects: Unified Analysis via H-likelihood*, Chapman and Hall.

Other readings: These include published or submitted journal articles, which are available on the course web page with one exception as noted:

Cui Y, Hodges JS, Kong X, Carlin BP (2010). Partitioning degrees of freedom in hierarchical and other richly-parameterized models. *Technometrics*, 52:124-136.

Hodges JS (1998). Some algebra and geometry for hierarchical models, applied to diagnostics (with discussion). *Journal of the Royal Statistical Society, Series B*, 60:497-536.

Hodges JS, Clayton MK (2011). Random effects old and new. Unpublished manuscript.

Hodges JS, Cui Y, Sargent DJ, Carlin BP (2007). Smoothing balanced, single-error-term analysis of variance. *Technometrics*, 49:12-25.

Hodges JS, Reich BJ (2010). Adding spatially-correlated errors can mess up the fixed effect you love. *The American Statistician*, 64:325-334.

Martinez-Beneito MA, Hodges JS, Marí-Dell-Olmo M (pre-publication version, 2014). Smoothed ANOVA Modeling. *CRC Handbook of Spatial Epidemiology*, eds. Andrew B Lawson, Sudipto Banerjee, Robert Haining, Lola Ugarte, to appear.

Reich BJ, Hodges JS (2008). Identification of the variance components in the general two-variance model. *Journal of Statistical Planning and Inference*, 138:1592-1604.

Reich BJ, Hodges JS, Carlin BP (2007). Spatial analyses of periodontal data using conditionally autoregressive priors having two classes of neighbor relations. *Journal of the American Statistical Association*, 102:44-55.

Reich BJ, Hodges JS, Zadnik V (2006). Effects of residual smoothing on the posterior of the fixed effects in disease-mapping models. *Biometrics*, 62:1197-1206.

Available from the instructor in hard-copy only.

Lavine M, Hochman D, Haglund D (2009). Optical images of the brain during surgery: Detection of hemodynamics in response to electrical stimulus. Unpublished manuscript.

COURSE OUTLINE/WEEKLY SCHEDULE

Below, H2014 refers to Hodges (2014), "RWC" refers to Ruppert, Wand, and Carroll (2003), LNP refers to Lee, Nelder, and Pawitan (2006), and RH refers to Rue and Held (2005), each listed above as a required or supplementary text. Journal articles are cited in standard format, e.g., "Reich et al (2006)", referring to the list of readings given above.

Weeks 1 through 7 are Part I of the course, covering syntaxes for richly parameterized models, focusing on the mixed linear model syntax and associated inferential tools. Weeks 8 through 14 are Part II of the course, showing the beginnings of a full theory of richly-parameterized models mostly by showing odd, mysterious, or incorrect results produced by standard analyses and developing theory to explain them. Homework assignments are concentrated in Part I of the course; during Part II, it is expected that students will work on their final class projects. The instructor will meet at least twice with each student regarding her/his final class project, or three times if enrollment is low enough to permit it.

Weeks 1, 2/ September 6, 11, 13. An introductory example intended to scare you, about how much anybody really understands even relatively simple models. The mixed linear model in the standard form, and the conventional (non-Bayesian) analysis of it. Bayesian analysis of the mixed linear model; pros and cons of non-Bayesian and Bayesian analyses, and problems with each approach.

Readings: H2014 Chapter 1 (sections 1.1-1.5), RWC chapters 1, 2 (through 2.6), 4 (except 4.9), 16.

1st homework assignment handed out on Wednesday 13 September, due Wednesday 20 September.

Week 3/ September 18, 20. An alternative formulation of the mixed linear model (the constraint-case formulation). Measures of complexity in a mixed linear model fit (degrees of freedom).

Readings: Required: H2014 Chapter 2; if you're keen on degrees of freedom, Cui et al (2010).

Week 4/ September 25, 27. Smoothing using penalized splines: bases, knots, penalized splines, rank of smoothers; penalized splines as mixed linear models: fitted values, inference, complexity of the fit (degrees of freedom).

Reading: H2014 Chapter 3; RWC Chapter 3, Chapter 4 section 9, Chapter 6 sections 1-4. Optional: RWC Chapter 6 sections 5-9.

2nd homework assignment handed out on Monday 25 September, due Monday 2 October.

Week 5/ October 2, 4. Additive models and models with interactions, represented as mixed models; using priors on degrees of freedom to control the extent of smoothing.

Reading: H2014 Chapter 4; RWC Chapters 7, 8, 9, and 12 (sections 1-3). If you're keen on smoothed ANOVA, Hodges et al (2007); if you're keen to see spatial smoothing in the context of smoothed ANOVA, Martinez-Beneito et al (2014).

3rd homework assignment handed out on Wednesday 4 October, due Wednesday 11 October.

Week 6/ October 9, 11. Spatial smoothing using mixed linear models: smoothing on a lattice using improper conditional autoregressive (CAR) models; two-dimensional splines, mostly with radial basis functions.

Reading: H2014 Chapter 5; RWC Chapter 13, sections 1-4.

4th homework assignment handed out on Wed 11 October, due Mon 23 October. This assignment is a proposal for the student's final class project; each student will meet individually with the instructor during week 8 to discuss the proposal.

Week 7/ October 16, 18. Time series using dynamic linear models (Kalman filters). Quick looks at the alternative syntaxes in RH and LNP. Summary of Part I of the course.

Reading: H2014 Chapters 6, 7. Optional: Browse RH or LNP.

Week 8/ October 23, 25. Introduction to Part II: What do we want in a theory of richly parameterized models? From linear models to richly-parameterized models: Mean structure. Generalizations of diagnostics from single-variance linear models, using the constraint-case formulation: Some diagnostics generalize, some don't, and we begin to see what's different about richly-parameterized models.

Reading: H2014 Chapter 8. If you'd enjoy seeing your professor embarrassed in print, read Hodges (1998) especially Section 5.2 and Wakefield's discussion.

Final class project: First meeting with the instructor about the final class project.

Week 9/ October 30, November 1. Continuing the previous week's material. Beginning: Collinearity and confounding in the presence of smoothing/shrinkage. Teaser: Weird things that happened fitting richly-parameterized models in four real applications. First weird thing: Adding a spatial random effect to a model makes an obvious association disappear (spatial confounding); mechanics of how it happens, how to interpret it, and what to do about it.

Reading: H2014 Chapter 9 (teasers), Section 10.1; if you're keen to read the original papers about spatial confounding, Reich et al (2006), Hodges & Reich (2010).

Final class project: Students are expected to be working on their final class projects.

Week 10/ November 6, 8. Continuing the previous week's material, including the Second weird thing.
Reading: Same as previous week's
Final class project: 2nd meeting with the instructor about the final class project.

Week 11/ November 13, 15. Third weird thing: Adding a random effect radically changes the fit of two other random effects (i.e., collinear or competing random effects). Generating and testing hypotheses about why this happens.

Reading: H2014, Chapter 12; Lavine et al (2009).
Final class project: Students are expected to be working on their final class projects.

Week 12/ November 20, 22. Traditional random effects (as in, e.g., Scheffé's classic ANOVA text) versus new- style random effects (i.e., most random effects discussed in this course), which have the mathematical form of a traditional random effect but which do not meet the traditional definition. Teaser for the last part of the course: Weird things related to estimation of unknowns in the variance structure.

Reading: Random effects old and new: H2014 Chapter 13. Teaser: Weird things ..., H2014 Chapter 14. Final class project: 3rd meeting with the instructor about the final class project.

Week 13/ November 27, 29. Beyond linear models: Variance structure. The general 2-variance model: Re- expressing the restricted likelihood as the likelihood from a particular generalized linear model, which enables many existing tools for examining how the data provide information about the two variances. Examples; application to some puzzles described in the previous week's Teasers.

Reading: H2014, Chapters 15, 16.
Final class project: Students are expected to be working on their final class projects.

Weeks 14,15/ December 4, 6, 11,13. Extending this approach to more complex models; two expedients of possible use when it can't be extended. Applications to more puzzles from the Teasers. A bit on estimates on the boundary of the parameter space and on restricted likelihoods and posterior distributions with multiple maxima. In-class presentation of student projects.

Reading: H2014, Chapters 17 (you can skip the gruesome details), 18, 19 (the latter two are short).

Final class project: Students will do a formal (ENAR-style) in-class presentation of their final class projects, and are expected to be working on the written version that's due during exam week.

The written version of the final class project is due Wednesday 20 December.

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

Grading will be A-F. 20% of the grade will be based on homework exercises, 60% will be based on the written version of the final class project; 20% will be based on the oral presentation of the final class project.

For additional information, please refer to the University's Uniform Grading Policy and Grading Rubric Resource at https://z.umn.edu/gradingpolicy **.**

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).

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 (https://plagiarism.iu.edu/certificationTests/).
Late Assignments	There will be no specific policies related to late assignments; at this advanced stage of your graduate career, and in a class as small as this one, it is expected that there will be few or none and that if you must be late on an assignment, you will make an arrangement with the instructor.
Attendance Requirements	You are expected to attend class. If you cannot make a class, e.g., because of illness, you are expected to tell the instructor because one person's absence has a notable effect on a class as small as this one will be.
Makeup Work for Legitimate Reasons	If you experience an extraordinary event that prevents you from completing coursework on time and you would like to make arrangements to make up your work, contact your instructor within 24 hours of the missed deadline if an event could not have been anticipated and at least 48 hours prior if it is anticipated. University policy recognizes that there are a variety of legitimate circumstances in which students will miss coursework, and that accommodations for makeup work will be made. This policy applies to all course requirements, including any final examination. Students are responsible for planning their schedules to avoid excessive conflicts with course requirements. 1. Instructors may not penalize students for absence during the academic term due to the following unavoidable or legitimate circumstances: illness, physical or mental, of the student or a student's dependent; medical conditions related to pregnancy; participation in intercollegiate athletic events; subpoenas; jury duty; military service; bereavement, including travel related to bereavement; religious observances; participation in formal University system governance, including the University Senate, Student Senate, and Board of Regents meetings, by students selected as representatives to those bodies; and activities sponsored by the University if identified by the senior academic officer for the campus or the officer's designee as the basis for excused absences. 2. Voting in a regional, state, or national election is not an unavoidable or legitimate absence. 3. Instructors are expected to accommodate students who wish to participate in party caucuses, pursuant to Board of Regents resolution (see December 2005 Board of Regents Minutes, p 147.) 4. For circumstances not listed in (1), the instructor has primary responsibility to decide on a case-by-case basis if an absence is due to unavoidable or legitimate circumstances and grant a request for makeup work. Because this course is entirely online and all materials are available to students from the first day of the term, we expect students to plan accordingly if travels or access to internet will cause them to miss a deadline. Note that our deadlines are generally set for 11:55 p.m. CST, so traveling to a different time zone will require additional planning. Further, circumstances that qualify for making up missed work will be handled by the instructor on a case-by-case basis; they will always be considered but not always granted. For complete information, view the U of M's policy on Makeup Work for Legitimate Absences (https://policy.umn.edu/education/makeupwork).
Extra Credit	There are no provisions for extra credit.