

Syllabus

Regression Methods for Correlated Data

STAT/BIOST 571, Winter 2017

Lectures: Mon/Wed/Fri 1:30-2:20 pm in Health Sciences **T-531 on Mon and Fri** and in UW Medical Center **RR-134 on Wed** (3 credits)

Instructor: Adam A. Szpiro (aszpiro@uw.edu (<mailto:aszpiro@uw.edu>)); office hours Wednesday 2:30-3:30 or by appointment in HSB F-673

TAs: Phuong Vu (phuongvu@uw.edu (<mailto:phuongvu@uw.edu>)), Li Wang (lxwang@uw.edu (<mailto:lxwang@uw.edu>)), and Brian Williamson (brianw26@uw.edu (<mailto:brianw26@uw.edu>)). TA office hours and locations are listed on the calendar.

Stat/Biost 571 covers analysis of correlated outcomes. It extends the regression methods for independent outcomes developed in Stat/Biost 570, which in turn build on the introduction to inference given in Stat/Biost 533 and earlier courses. In 571, we will discuss the motivation, interpretation and implementation of various techniques used when independence does not hold.

This is a methods course, and the emphasis will be on application and interpretation of the methods discussed, in practical circumstances. Formal proofs and similarly-mathematical arguments are not a major part of the course, but you will be expected to be able to choose appropriate methods, implement them accurately, interpret the output accurately, and explain output in language appropriate for scientific collaboration. To do this, statistical insight, mathematical and computational skill, and ability to communicate are all essential.

Prerequisites:

- *STAT/BIOST 570 Analysis of independent outcomes*: in particular, you are assumed to have experience accurately interpreting regression coefficients, making inference based on multivariate linear and non-linear regressions, and identifying and using non-parametric, semi-parametric, and parametric forms of inference.
- *STAT 512/513 Probability and Statistical Inference*, or equivalent: in particular, you are assumed to be familiar with parametric likelihood-based and Bayesian statements, as used with simple regression-based models.
- Familiarity with R. Material up to the level seen in 570 should be sufficient: in particular, you should be familiar with R's formula syntax, and should be able to design and implement simulation studies.

Learning objectives:

By the end of the course the student should be able to

- Explain why conventional regression models (as reviewed in 570) are inappropriate for correlated data, and explain the likely detrimental effects of their use in the correlated context.
- Describe the key differences between the generalized estimating equations and likelihood approaches to modeling correlated data. In particular the student should be able to describe the differences between the conditional and marginal approaches to modeling, and the advantages and drawbacks of each.
- Suggest appropriate approach(es) to modeling correlated outcomes, critically evaluate the fit of the fitted models, and interpret the estimated regression coefficients.
- Outline implementation strategies for generalized estimating equations and likelihood approaches, including Bayesian methods.

Required textbooks (on reserve at Health Sciences Library):

Diggle, Heagerty, Liang, and Zeger (2013). *Analysis of Longitudinal Data (second edition)*

Wakefield (2013). *Bayesian and Frequentist Regression Methods*

Software:

We will primarily use the R statistical software packages. R is a free software package available on Windows, MAC OS X, and Linux operating systems. It can be downloaded from the following website: <http://cran.r-project.org/>.

Assessment:

This is a graded course; the "pass" level is 3.0. Your grade will be based on:

- Homework (20%): there will be assignments most weeks, distributed online on Wednesday and due at the beginning of lecture the following Wednesday. In weeks with midterm exams there will be either no homework or shorter assignments.
- Class participation (10%): it helps the whole class if you ask questions during lectures, or answer questions posed by the instructor
- Two midterm exam (20% each): these will be one hour in-class closed book exam during a regular class period.
- Final exam (30%): this will be a two hour in-class closed book exam during finals week (tentatively scheduled for Monday, March 13 from 2:30 pm - 4:20 pm, room TBD)

Please contact the instructor **ahead of time** if (for some good reason) you are not able to meet homework and/or exam deadlines; extensions may be granted, but only prior to deadlines. Completing all homework assignments and exams is strongly recommended. In this course, doing the homework exercises and participating in class are considered integral to both learning the material, and acquiring skills in critical reasoning.

Discussion of **homework exercises** is encouraged, including asking the instructor or TAs for clarification about what you are expected to do. However, you must **write and code your homework on your own**. Exams will be closed-book and must be completed **entirely on your own**. Please review the **University Policy** [↗ \(http://depts.washington.edu/grading/pdf/AcademicResponsibility.pdf\)](http://depts.washington.edu/grading/pdf/AcademicResponsibility.pdf) if any aspect of this is unclear.

- Homework must be word processed; using LaTeX is recommended, and is good preparation for 572 and future research.
- Printed copies of the assignments are to be turned in at the beginning of lecture on the due date.
- Producing publication-ready write-ups is not expected, nor it is likely to be a good use of your time. Hand-written annotations of text/graphics are acceptable, but please write clearly.
- All homework and exams must be written up in full sentences and paragraphs. Handing in raw mathematical arguments, or raw computer code or computer output will not receive credit.
- In addition to your write-up, please include annotated R code for all problems in an appendix. This will be increasingly important as the homework assignments start to involve more complicated calculations.
- Some questions may ask you to write appropriately for a co-author who is not a statistician; if this is unfamiliar to you please practice, by explaining your work to a non-statistician friend.

In-class exams will be hand-written, but please note that neatness and legibility of your responses are required for full credit. You are responsible for material covered in lectures as well as understanding that can be gleaned from homework assignments (your own work and the keys).

The course **discussion board** is the primary forum for asking questions electronically about course material or homework. Other students will often provide helpful feedback, and the board will also be monitored by TAs and the Instructor. For questions that are not appropriate for the discussion board, please direct **email** to the Instructor and all three TAs.

Access and Accommodations:

Your experience in this class is important to me. If you have already established accommodations with Disability Resources for Students (DRS), please communicate your approved accommodations to me at your earliest convenience so we can discuss your needs in this course.

If you have not yet established services through DRS, but have a temporary health condition or permanent disability that requires accommodations (conditions include but not limited to; mental health, attention-related, learning, vision, hearing, physical or health impacts), you are welcome to contact DRS at 206-543-8924 or [uwdrs@uw.edu \(mailto:uwdrs@uw.edu\)](mailto:uwdrs@uw.edu) or [disability.uw.edu. ↗ \(http://depts.washington.edu/uwdrs/\)](http://depts.washington.edu/uwdrs/) DRS offers resources and coordinates reasonable accommodations for students with disabilities and/or temporary health conditions. Reasonable accommodations are established through an interactive process between you, your instructor(s) and DRS. It is the policy and practice of the University of Washington to create inclusive and accessible learning environments consistent with federal and state law.

Academic Integrity:

Students at the University of Washington (UW) are expected to maintain the highest standards of academic conduct, professional honesty, and personal integrity.

The UW School of Public Health (SPH) is committed to upholding standards of academic integrity consistent with the academic and professional communities of which it is a part. Plagiarism, cheating, and other misconduct are serious violations of the University of Washington **Student Conduct Code** [↗ \(http://www.washington.edu/cssc/student-conduct-overview/student-code-of-conduct/\)](http://www.washington.edu/cssc/student-conduct-overview/student-code-of-conduct/) (WAC 478-120). We expect you to know and follow the university's policies on cheating and plagiarism, and the **SPH Academic Integrity Policy** [↗ \(http://sph.washington.edu/students/academicintegrity/\)](http://sph.washington.edu/students/academicintegrity/). Any suspected cases of academic misconduct will be handled according to University of Washington regulations. For more information, see the University of Washington **Community Standards and Student Conduct** [↗ \(http://www.washington.edu/cssc/\)](http://www.washington.edu/cssc/) website.

Teaching Assistants:

If you have any concerns about the class or your TA, please see the TA about these concerns as soon as possible. If you are not comfortable talking with the TA or not satisfied with the response that you receive, you may contact the Department of Biostatistics Associate Director of Academic Affairs ([biostgp@uw.edu \(mailto:biostgp@uw.edu\)](mailto:biostgp@uw.edu)). If you are still not satisfied with the response that you receive, you may contact the Department

of Biostatistics Chair (bchair@uw.edu (<mailto:bchair@uw.edu>)). You may also contact the Graduate School at G-1 Communications Building, by phone at 206-543-5139 or by email at raan@uw.edu (<mailto:raan@uw.edu>).