2021 Advanced Regression Methods for Independent Data

BIOSTAT/STAT 570

Jon Wakefield

Departments of Statistics and Biostatistics
University of Washington
jonno@uw.edu

Logistics

- **Lectures:** Monday/Wednesday/Friday 1.30–2.20 (Thomson 125).
- **Coursework:** Weekly. Take-home midterm and final exams.
- **Grading:** Coursework (50%), mid-term (25%), final (25%).
- **JW Office Hour:** Monday 3:30-4:30, Department of Biostatistics, Room 331.1.C18, Hans Rosling Center for Population Health. Or by appointment (Email: `jonno@uw.edu`).
- **TAs:**
  - Rahul Biswas (Department of Statistics), `rbiwas1@uw.edu`
  - Anna Neufeld (Department of Statistics), `aneufeld@uw.edu`
- **TA Office Hours:**
  - To be arranged.
  - To be arranged.
- **Computing** will be carried out using `R`.
- **Next quarter:** Stat/Biostat 571 Advanced Regression Methods for Dependent Data.

Main Text

Wakefield, J.C. (2013). *Bayesian and Frequentist Regression Analysis*. Springer. e-book available from library.

- Chapters 1–3, 5–7 are covered in Biostat 570.
- Chapters 8–9 in Biostat 571.

You will find `R` code for all examples, and typos (!) at:

`http://faculty.washington.edu/jonno/regression-methods.html`

It would be advantageous for you to read the relevant book chapters before lectures.

Supplementary Texts

- Davison, A.C. (2003). *Statistical Models*. Cambridge University Press.
- McCullagh, P. and Nelder, J.A. (1989) *Generalized Linear Models, Second Edition*, Chapman and Hall.
- Cameron, A.C. and Trivedi, P.K. (1998). *Regression Analysis of Count Data*, Cambridge University Press.
- Gelman, A, Carlin, J.B., Stern, H., D. Dunson, A. Vehtari and Rubin, D.R. (2013). *Bayesian Data Analysis, Third Edition*, CRC Press.

Course Outline

Chapter 1: Overview and Motivating Examples
Steps and context of data analysis; Nature of Randomness; Types of data; Implications for inference; Historical context.

Chapter 2: Frequentist Inference
Justification of approach; Methodology; Likelihood; Quasi-likelihood.

Chapter 3: Bayesian Inference
Justification of approach; Methodology; Prior choice; Computation.

Chapter 5: Linear Models
Least squares; Likelihood and Bayesian approaches; Interpretation; Bias and Confounding; Diagnostics.

Chapter 6: General Regression Models
Generalized linear models; Nonlinear models.

Chapter 7: Binary Data Models
Simpson's paradox; Logistic regression; Bias and Confounding.

What's this course all about?

We are not going to talk about **prediction** but instead about **inference for parameters**.

Learning about different philosophical approaches/methods/model classes, and ways to assess the appropriateness of models.

When you read applied/methods papers I want you to be able to answer important questions such as:

- What is the inferential paradigm?
- What are the key assumptions?

Computation is an additional issue.

While we should fit models supported by science, some models are better than others in terms of statistical properties; in particular, are key summaries consistently estimated under model misspecification?