AARHUS UNIVERSITY

Methods 2 - 1

Chris Mathys

BSc Programme in Cognitive Science

Spring 2023

Methods 2 – Overview

- The General Linear Model (GLM)
- Regression modelling
- Mathematical foundations
  - Linear algebra (vectors, matrices, determinants, eigen-analysis,…)
  - Calculus (infinite series, derivatives, integrals,…)
- Generalizations of the GLM (e.g., logistic regression)

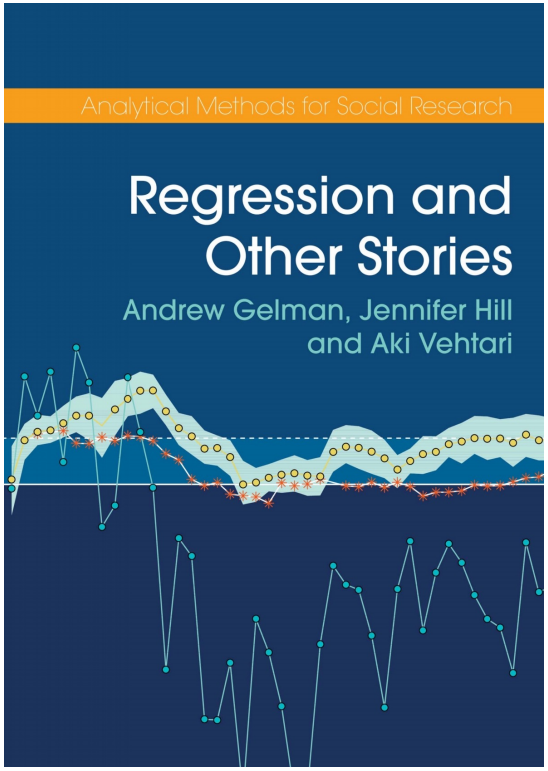
Resources

Textbook:

Gelman, A., Hill, J., & Vehtari, A. (2020). *Regression and Other Stories* (Analytical Methods for Social Research). Cambridge: Cambridge University Press. doi:10.1017/9781139161879

Please get a copy!

Free PDF: https://avehtari.github.io/ROS-Examples/

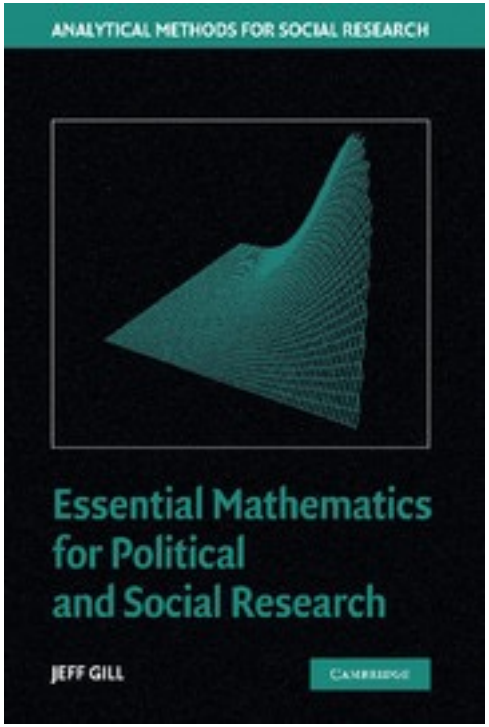


Resources

Textbook:

Gill, J. (2006). *Essential Mathematics for Political and Social Research* (Analytical Methods for Social Research). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511606656

No need to buy it! You have access to PDFs of the chapters via the Royal Library.



Schedule

Course week	Week of year	Topics and readings
1	6	Regression and the GLM: overview, data and measurement, (GHV[1] 1,2)
2	7	Basic methods, statistical inference (GHV 3,4)
3	8	Statistical inference (continued), simulation (GHV 4,5)
4	9	Math basics: functions, equations, polynomials, logarithms (Gill[2] 1)
5	10	Linear algebra basics: vectors, matrices, norms, transposition (Gill 3)
6	11	More linear algebra: geometry, determinants, rank, inversion, eigenvectors (Gill 4)
7	12	Scalar calculus: derivatives, integrals, fundamental theorem (Gill 5)
8	13	More calculus: root finding, extrema, Lagrange multipliers, vector calculus (Gill 6)
9	15	Conceptual foundations and history of the GLM, model fitting (GHV 6,7,8)
10	16	Fitting GLMs: prediction, Bayesian inference (GHV 9)
11	17	Multiple predictors, interactions (GHV 10)
12	18	Model comparison, assumptions and diagnostics (GHV 11)
13	19	Transformations, predictive simulations (GHV 12) [no class, just lecture]
13a	21	Final class to wrap up portfolio writing

[1] Gelman, A., Hill, J., & Vehtari, A. (2020). *Regression and Other Stories* (Analytical Methods for Social Research). Cambridge: Cambridge University Press. doi:10.1017/9781139161879

[2] Gill, J. (2006). *Essential Mathematics for Political and Social Research* (Analytical Methods for Social Research). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511606656

Resources

Code:

- This course's repository: https://github.com/methods-2-f23/methods-2-resources

- All code and data in the book: https://github.com/avehtari/ROS-Examples

- **Please get a free GitHub account** – or wait until class tomorrow, we'll set everything up there

Resources

Videos:

- This course is on YouTube! https://www.youtube.com/playlist?list=PLvJwKACYy5_MTdnrzxx_1sN389dS9OB3S
- Order slightly different: we'll do GHV chapters 1-5 in the first 3 weeks, but after that, you can watch the videos in the order of the playlist

Exam

https://kursuskatalog.au.dk/en/course/115680/Methods-2-The-General-Linear-Model

"Ordinary examination and re-examination:
The exam consists of a portfolio containing a number of assignments. The total length of the portfolio is: 3-7 assignments.

Their form and length will be announced on Blackboard by the teacher at the start of the semester. The portfolio may include products. Depending on their length, and subject to the teacher's approval, these products can replace some of the standard pages in the portfolio.

It must be possible to carry out an individual assessment. So if some parts of the portfolio have been produced by a group, it must be stated clearly which parts each student is responsible for, and which parts the group as a whole is responsible for.

The complete portfolio must be submitted for assessment in the Digital Exam system. Each student submits a portfolio."

Exam

- Portfolio consisting of 3 assignments
- Each assignment will require you to create an R Markdown notebook consisting of a mix of text and code.
- Due

  1. End of week 9 (Sunday 5 March, 23:59)
  2. End of week 13 (Sunday 2 April, 23:59)
  3. End of week 17 (Sunday 30 April, 23:59)

You will receive a (short) feedback message from us on your portfolio assignments that you can use for improvements before finalizing your hand-ins.

Regression

- What is regression? Ideas, suggestions?

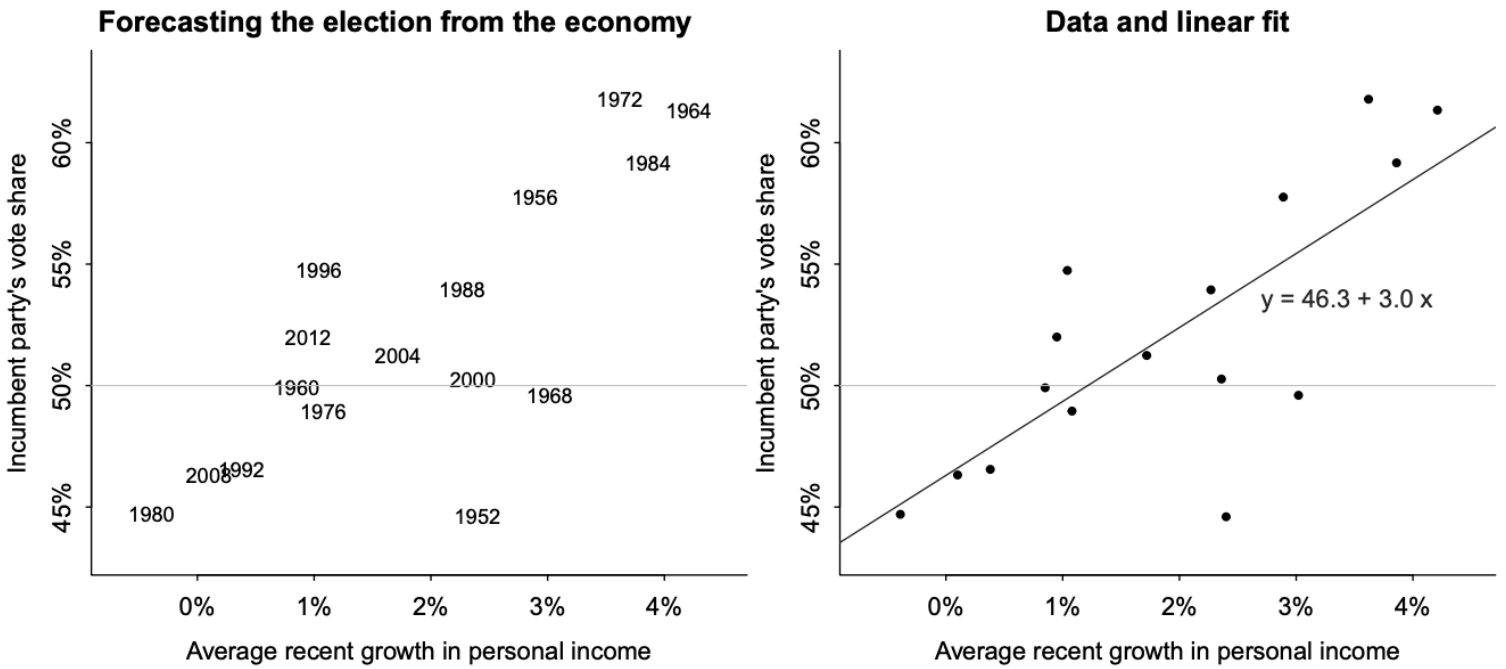


Figure 1.1: *Predicting elections from the economy: (a) the data, (b) the linear fit,* $y = 46.3 + 3.0x$.

Linearity

- When is something linear in mathematics?

Linear regression

- If we combine regression and linearity, what do we get?

The general linear model (GLM)

- What is general about it?

This is part of statistics. But what about statistics?

Three challenges

1. Generalizing from sample to population
2. Generalizing from treatment to control group
3. Generalizing from observed measurements to underlying constructs of interest

Challenges in regression: examples

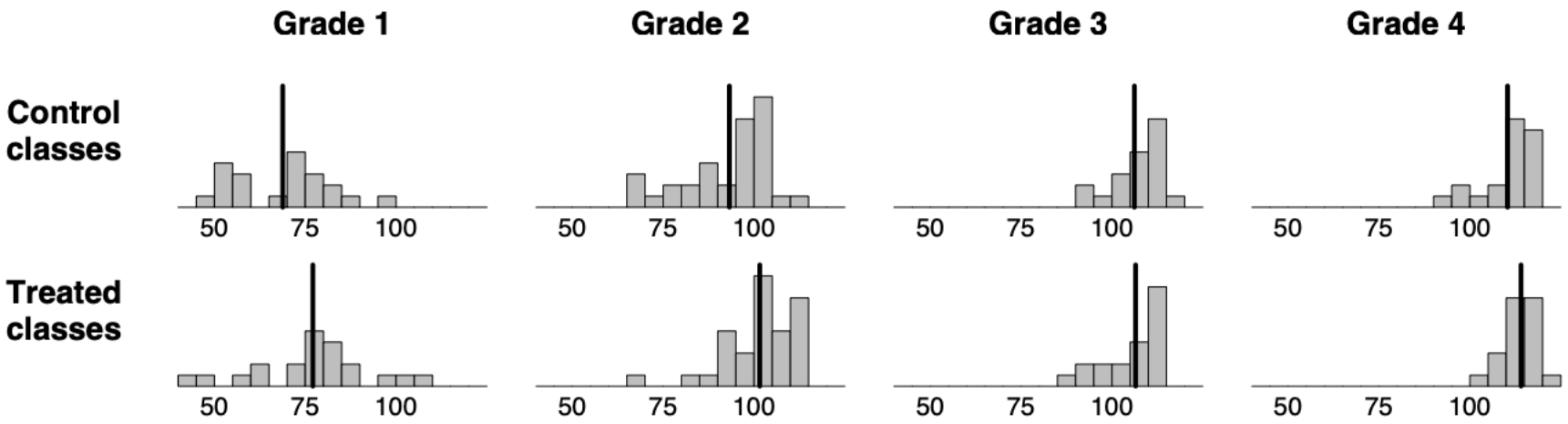


Figure 1.2 *Post-treatment classroom-average test scores from an experiment measuring the effect of an educational television program, The Electric Company, on children's reading abilities. The dark vertical line in each histogram shows the average for the corresponding group of classrooms.*

Challenges in regression: examples

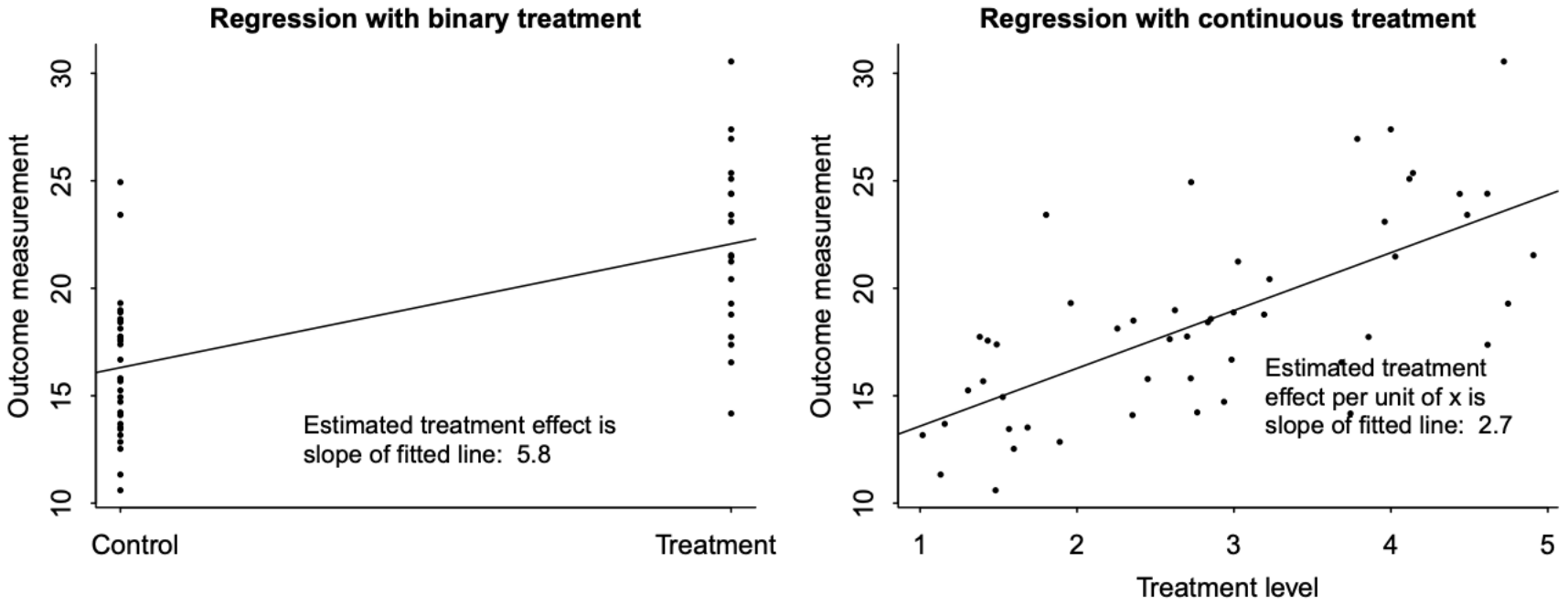


Figure 1.5 *Regression to estimate a causal effect with (a) simple comparison of treatment and control, or (b) a range of treatment levels.*

Challenges in regression: examples

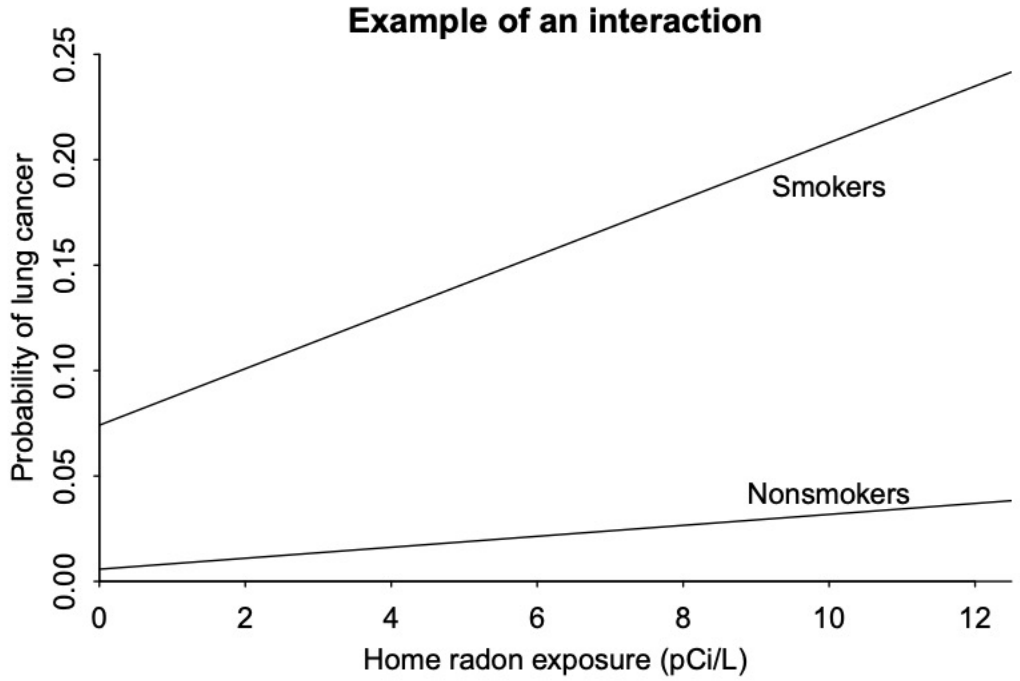


Figure 1.7 *Lifetime added risk of lung cancer for men, as a function of average radon exposure in picocuries per liter (pCi/L). The relation between cancer rate and radon is different for smokers and nonsmokers.*

Challenges in regression: examples

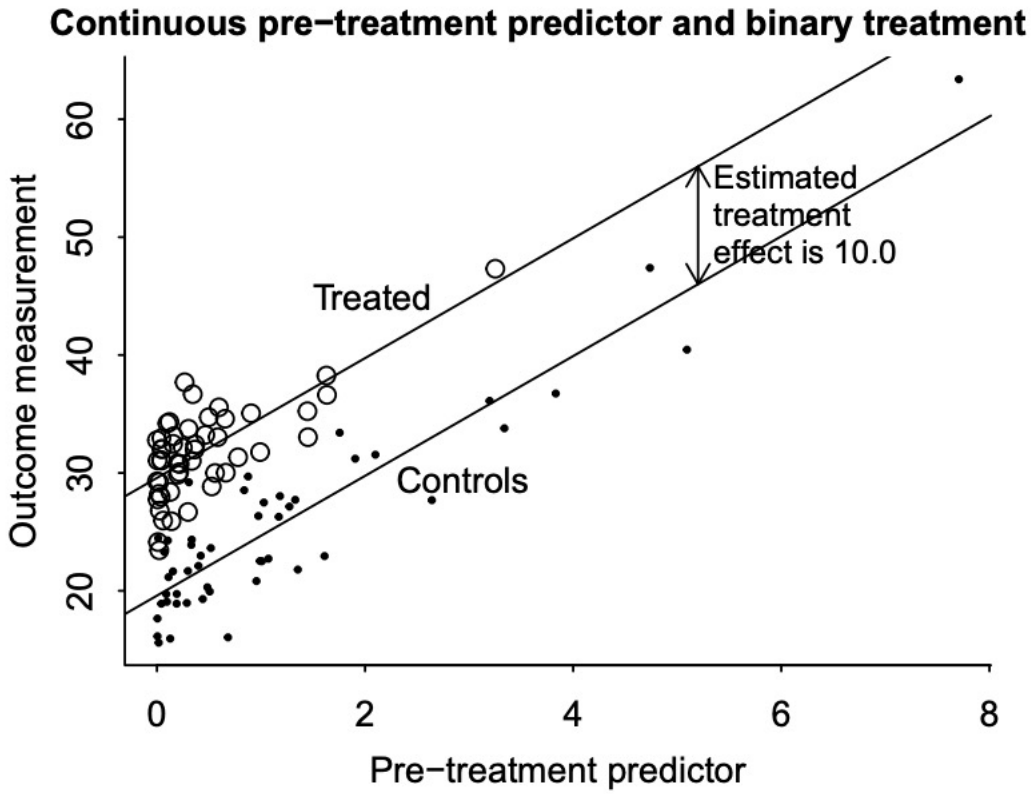


Figure 1.8 *Hypothetical data with a binary treatment and a continuous pre-treatment variable. Treated units are displayed with circles on the scatterplot, and controls are shown with dots. Overlaid is a fitted regression predicting the outcome given treatment and background variable, with the estimated treatment effect being the difference between the two lines.*