STAT 5412
Biostats 2: Generalized Linear Models

Instructor: Justin French TA: Caleb Hughes

January 17, 2024

RAS 128
10:00 AM - 11:50 AM, M,W,F
Email (preferred): justin.french@sulross.edu
Phone: (432) 837-8505
Office Hours: Fridays, 1:00 PM to 4:00 PM, or by appointment.

1 Course Overview

This class continues the basic concepts, mathematics, and applications of modern statistical methods. We will pick up with the general linear model, and steadily build towards framing ecological hypotheses as statistical models. Then we will learn to assess the performance of those models and explore the limitations of simplistic assumptions. We will move beyond the foibles of statistics past and work towards a general statistical literacy throughout the course. By the end of this course you will be able to:

1. Summarize the expected value and nature of uncertainty in your data.
2. Pose a hypothesis in terms of a statistical model.
3. Fit (generalized) linear (mixed) models to real data.
4. Interpret the estimated parameters of your models.
5. Assess the performance of fitted models.

2 Required Texts

Fieberg, J. 2022. Statistics for Ecologists. An open-source online textbook.
https://fw8051statistics4ecologists.netlify.app/

3 General Assignments

I will assign exercises each week to reinforce the material. All work in this class will be done in R, which is a free statistical analysis package (and so much more, as you will see). It is best that you install both R and R Studio, which is a convenient interface. They are available for free download at:

- **R:** https://cloud.r-project.org/
- **R Studio:** https://rstudio.com/products/rstudio/download/

All assignments integrate written work and statistical results, so we will utilize a typesetting program called LaTeX within R Studio to do all assignments. The easiest and most reliable way to get started with LaTeXin R Studio is called TinyTex, which can be installed following the procedures shown here: https://www.youtube.com/watch?v=oF_KI4SLuBs&ab_channel=AmeliaMcNamara

4 Semester Project

By the end of the semester you will complete a full-blown analysis of a complex data set using what you have learned. The final week of class will consist of presentations where you present your findings as you would at a professional conference. In general your presentation should include the following:

1. Frame your research question, as well as it's motivations and potential implications
2. Discuss data collection and the model you will use to answer the question
3. Implement your analysis
4. Interpret results
5. Make inferences and discuss their implications

Because many of you are still collecting data, it may not be possible to produce final results within the semester. However, you will have laid the foundation of your analysis from top to bottom in a way that only requires adding in new data once it is available and re-interpreting results.

5 Exams

Exams will be implemented in R using LaTeX. You will receive the exam via email on the date it is scheduled. You will have 24 hours to complete it and return it to me. Late submissions will not be accepted and will result in a 0.

6 Grading Policy

Grades are based on weekly assignments, 3 exams, and the semester project. Each of these categories contribute contribute 33.3% of your final grade.

Letter grades follow: $100 > A \geq 90 > B \geq 80 > C \geq 70 > D \geq 60 > F$. There is no curve.

Late work will be eligible for up to 90% credit if 1 day late, 50% credit if 2 days late, and will not be graded if 3 or more days late, resulting in a 0. This policy will be strictly enforced.

7 Attendance

Showing up is the only way to get the material you need. If you don't come to class, your grade will reflect it with no penalty needed from me. In the event of an excused absence, make arrangements with me to go over material ahead of time.

8 Academic Dishonesty

Academic dishonesty includes copying, sharing, or obtaining information from an unauthorized source, attempting to take credit for the intellectual work of another person, falsifying information, and giving or receiving information about a test, quiz, or assignment to other students. Any student involved in academic dishonesty will receive no credit (0) for work done and/or may be penalized in accordance with published University Rules.

9 Counseling and Accessibility Services

Sul Ross State University is committed to equal access in compliance with the Americans with Disabilities Act of 1973. It is the student's responsibility to initiate a request for accessibility services. Students seeking accessibility services must contact Mary Schwartze, M. Ed., L.P.C., in Counseling and Accessibility Services, Ferguson Hall, Room 112. The mailing address is P.O. Box C-122, Sul Ross State University, Alpine, Texas 79832. Telephone: 432-837-8691. E-mail: mschwartze@sulross.edu.

10 Course Schedule (Tentative)

- Module 1: Generalizing the Linear Model
  - **Week 1 (1/15-1/19):** Beyond the General Linear Model
    - Lecture 1: MLK Day (No Class)
    - Lecture 2: Almost linear: transformations
    - Lecture 3: Undecipherable meanings: Foibles of statistics past
    - Reading: Arcsine is Asinine
    - Exercise: Bending linear models to your will
  - **Week 2 (1/22-1/26):** Droves of Distributions 1
    - Lecture 1: A distribution for everything: Exponential families
    - Lecture 2: Discrete distributions
    - Lecture 3: Discrete distributions, Con't.
    - Reading: TBD
    - Exercise: A primer on discrete probability distributions
  - **Week 3 (1/29-2/2):** Droves of Distributions 2
    - Lecture 1: Continuous events revisited
    - Lecture 2: Continuous distributions
    - Lecture 3: Continuous distributions, Con't.
    - Reading: TBD
    - Exercise: A primer on continuous probability distributions
  - **Week 4 (2/5-2/9):** Link Functions
    - Lecture 1: A different kind of transformation
    - Lecture 2: Interpreting beta coefficients, again
    - Lecture 3: Models as hypotheses in GLMs
    - Reading: TBD
    - Exercise: Logistic, Poisson, and Gamma regression
  - **Week 5 (2/12-2/16):** Maximum Likelihood
    - Lecture 1: Probability density functions and a simple MLE
    - Lecture 2: MLE and covariates
    - Lecture 3: What you really need: How to solve your own problems
    - Reading: TBD
    - Exercise: Maximize your own likelihood: An identifiability crisis

- **Week 6 (2/19-2/23):** Likelihood and Model Selection
  - Lecture 1: MLE, Information, and Entropy: Practical approaches to model selection
  - Lecture 2: TCTWS Conference
  - Lecture 3: TCTWS Conference
  - Reading: TBD
  - Exercise:

- **Week 7 (2/26-3/1):** A Framework for Model Construction
  - Lecture 1: Laying out some general steps
  - Lecture 2: Discuss class projects
  - Lecture 3: Exam Day!!!
  - Reading: TBD
  - Exercise:

- Module 2: Mixing it Up: Random Effects and Hierarchy
  - **Week 8 (3/4-3/8):** Basics of Mixed Models 1
    - Lecture 1: Everyone gets an average! - Random Intercepts
    - Lecture 2: Variable relationships - Random Slopes
    - Lecture 3: Building models with your new Legos
    - Reading: TBD
    - Exercise: Fitting and interpreting a simple GLMM

  - **Spring Break (3/11-3/15):** No Class.
  - **Week 9 (3/18-3/22):** Basics of Mixed Models 2
    - Lecture 1: Matching models to study design
    - Lecture 2: Model diagnostics with mixed effects
    - Lecture 3: Model selection with mixed effects
    - Reading: TBD
    - Exercise: Understanding a GLMM through simulation

  - **Week 10 (3/25-3/29):** Intermediate Mixed Models
    - Lecture 1: Multilevel models
    - Lecture 2: Sophisticated hypotheses as hierarchical models
    - Lecture 3: Limits of MLE at finite sample sizes
    - Reading: TBD
    - Exercise: How much data does it take?

  - **Week 11 (4/1-4/5)**: A Primer on Bayesian Inference
    * Lecture 1: Bayes Theorem, Likelihood, and Prior information
    * Lecture 2: Bayesian thinking
    * Lecture 3: Fitting a hierarchical model in JAGS
    * Reading: TBD
    * Exercise: It's probably this: change my mind

  - **Week 12 (4/8-4/12)**: Internalizing the Debate
    * Lecture 1: Really thinking about probability
    * Lecture 2: Reviewing *everything*
    * Lecture 3: Exam day!
    * Reading: TBD
    * Exercise:

- Module 3: Ecological Applications

  - **Week 13 (4/15-4/19):** Ecological Applications
    * Lecture 1: Survival Models
    * Lecture 2: Time- and Space-To-Event Models
    * Lecture 3: State-Space Models
    * Exercise:

  - **Week 14 (4/22-4/26):** Ecological Applications
    * Lecture 1: Distance Sampling
    * Lecture 2: Occupancy Modelling
    * Lecture 3: What lies beyond...
    * Exercise:

  - **Week 15 (4/29-5/3):** Show me what you've learned!
    * Lecture 1: Project Presentations
    * Lecture 2: Project Presentations
    * Lecture 3: None - Finals Begin

  - **Week 16 (5/6-5/8):** Final Exam Week
    * No Final, you're done!