Intro to GLMs Group Exercise

Answer the questions below and submit your answers on Gradescope by 3:45PM.

1. Comprehension questions

- In logistic regression, which distribution is assumed for the outcome Y? What is the mean of this distribution?

- What is a link function?

- Why do we use link functions?

- Why do we use log odds instead of directly modeling Y or p?

- How do we define odds?

2. Practice problem (you don't need to use code)

A study is conducted to assess the effectiveness of a new drug to treat back pain. 735 participants first rate their baseline pain level. 366 are then assigned to the experimental drug and everyone else is assigned to the placebo group (participants do not know which group they are in). After 2 weeks, participants rate their level of back pain again. Responses are compared to initial ratings and participants are grouped into either "decreased pain" or "increase/no change." 482 participants experienced decreased pain, of which 289 were in the experimental drug group.

- Create the 2x2 table for this problem

- Calculate the probability of decreased pain in the experimental drug group and in the placebo group

- Calculate the odds ratio of decreased pain for the experimental group compared to the placebo group

- Write a sentence to interpret the odds ratio in the context of the problem

- If we were to conduct a hypothesis test to assess statistical significance of the odds ratio, what should the null value be?

3. If you have extra time, write your own example of a 2x2 table relevant to your domain of interest.