Quiz 5

Name:

Instructions: You have 15 minutes to complete this quiz. You may use any notes or resources, including your laptop and the internet (R documentation, StackOverflow, Wikipedia, etc.). You may not communicate with anyone else during the quiz.

There are four short questions, and one bonus question. Relax and do your best! Try to answer as much as possible.

The data is described on page 2, and the scenario is described on page 3. The questions are on pages 4 – 6.

Data

What is the relationship between drinking alcohol and an overdrawn bank account for college students? Researchers hypothesize that drinking large quantities of alcohol and overdrawing a bank account both indicate risky behavior, and so they suspect a positive association.

To investigate this question, researchers collected data on a random sample of 450 undergraduates at Mississippi State University and the University of Mississippi (two large universities). Each row in the data represents one student, and we will focus on two variables:

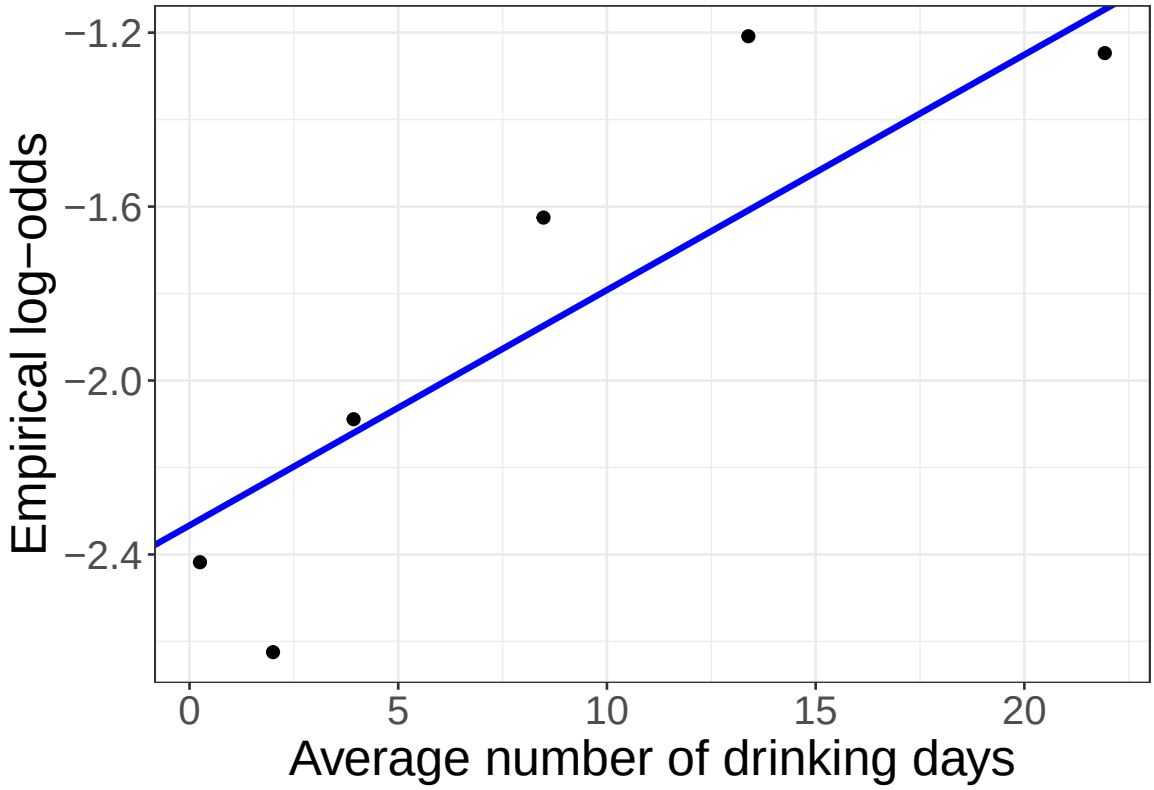
- `DaysDrink`: the number of days drinking alcohol, in the past 30 days
- `Overdrawn`: whether the student has ever overdrawn a bank account (0 = no, 1 = yes)

Scenario

The fitted logistic regression model is

$$\log\left(\frac{\hat{\pi}}{1-\hat{\pi}}\right) = -2.21 + 0.047 \text{ DaysDrink},$$

where $\pi = P(\text{Overdrawn} = 1)$. To assess the logistic regression model, the researchers create an empirical log-odds plot:



Questions

Question 1: What are the estimated odds of having overdrawn a bank account for a student who drank no alcohol in the last 30 days?

Question 2: What is the estimated change in odds associated with an increase of 1 day of drinking in the last 30 days?

Question 3: What is the estimated probability of having overdrawn a bank account for a student who drank alcohol on 10 of the last 30 days?

Question 4: Assess the shape, independence, and randomness assumptions for the researchers' logistic regression model.

Bonus question: The logistic regression model assumes that the change in odds associated with an increase of 1 drinking day is constant (as in Question 2). Is the change in *probability* of being overdrawn also constant? Explain your reasoning.