Homework 7 – Stat 230 – Fall 2022

Due date: Friday, October 28

Complete the following exercises and submit your assignment via gradescope (linked on the course webpage).

> **Note**
>
> When the book asks you to "Interpret the model in terms of the odds ratio," it is asking you to interpret the slope of the logistic regression model.

Problems to start after class Oct 21

Q1

The file `medschool.csv` contains data on 55 medical school applicants from a liberal arts college in the Midwest. We will focus on two variables, `Acceptance` (1 = accepted to medical school, 0 = not accepted) and `MCAT`, the student's score on the MCAT exam.

You can load the data using the command:

```
medschool <- read.csv("http://aloy.rbind.io/data/medschool.csv")
```

(a) Fit the logistic regression of acceptance status on MCAT score. Report the fitted logistic regression model ($\text{logit}(\hat{\pi}) = \cdots$).

(b) For somebody who scored 30 on the MCAT, find the probability they were accepted.

(c) For somebody who scored 30 on the MCAT, find the odds of being accepted (to not being accepted).

(d) Compare the odds of acceptance for somebody who scored a 35 to somebody who scored a 30 on the MCAT and give a sentence interpreting this number.

(e) Interpret the coefficient of MCAT in your model (in terms of odds).

Q2

Chapter 7 exercise E.1

- After class on October 21, you will be able to complete part (a)
- After class on October 24, you will be able to complete parts (b)-(c)

Use the following code to load the data. Notice that `*` is used to denote missing values, so we add the `na.strings` argument to handle this.

```
birdnest <- read.csv("https://aloy.rbind.io/kuiper_data/Birdnest.csv", na.strings = "*")
```

Q3

Chapter 7 exercise E.2

- After class on October 21, you will be able to complete part (a)
- After class on October 24, you will be able to complete parts (b)-(c)

```
donner <- read.csv("https://aloy.rbind.io/kuiper_data/Donner.csv")
```

Problems to start after class Oct 24

Q4

Chapter 7 exercise E.5