STAT 213 (SUMMER 2021): HW6

DUE ELECTRONICALLY BY FRIDAY 8/06/21

INSTRUCTIONS AND TECHNICAL TIPS

Write up your solutions and save/upload the file(s) to the RStudio server in the folder `stat213/turnin/hw6/` inside your Home directory. Include `hw6` (all lowercase) in your filenames.

Using an RMarkdown document is recommended, but not required.

RMarkdown Format. If you do use RMarkdown, R code should go in code chunks, and verbal commentary (and any math equations) should go outside code chunks.

Use section headings (a line starting with one or more `#` symbols) to demarcate the start of a problem. You can nest headings by using an additional `#` symbol for each level of nesting: one for a top-level heading, two for the next level within that, etc.

Periodically "Knit" your file to verify that it is working correctly.

"Knitting" Troubleshooting. If your code runs chunk by chunk but won't Knit, try clearing your environment (broom icon in the upper right) and running chunk by chunk from the start again. The most common cause is an undefined variable. This can happen if you change your variable names some places but not others, but a very common reason is that you read in the data from a file using a menu button instead of using the `read.file()` command, which means the "reading in the data" step is not recorded in your document. If you can't Knit, I won't be able to run your code either!

If you aren't able to Knit directly to `.pdf`, it may be because you are using special characters (such as $\neq$) in your `.Rmd`. Replace these with plain text and try again.

If you have done this and are still unable to Knit directly to `.pdf`, convert your Knitted `.html` or `.docx` into a `.pdf` (possibly by "printing" the `html` from your browser to a file) and save that there. A uniform file format across students will make grading 64 problem sets much more streamlined. Thanks!

What to turn in. If using Markdown, turn in the source file (ending in `.Rmd`) and the Knitted output file (preferably as `.pdf`). Otherwise, just turn in a `.pdf`.

Date: August 2, 2021.

Verifying receipt of your work. When I run the script that collects your work after the due date, you will see a "receipt" file in the ~/stat213/receipts/hw6/ directory. If this does not appear within 24 hours or so of you submitting your assignment, let me know.

PROBLEMS

1. What can go wrong if we try to use linear regression with no changes to model the probability of a particular outcome for a binary response variable? Give one problem related to the **predictions** such a model gives, and another related to **statistical inference** using such a model.

2. Convert the following **probabilities** to **odds**: (a) 0.5, (b) 0.9, (c) 0.1.

3. Convert the following **odds** to **probabilities**: (a) 2:1 (or 2), (b) 10:1 (or 10), (c) 1:4 (or 0.25).

4. If the **probability** of recovery with treatment is 0.3 and the probability of recovery without treatment is 0.1, find the **odds ratio** for recovery when treated vs not treated.

5. Suppose that we have a logistic model with $\beta_0 = 5$ and $\beta_1 = 2$. Explain what happens to a plot of the **probability form** of the model in each of the following circumstances.

   (a) β_1 decreases to 1

   (b) β_0 increases to 8

   (c) β_1 changes sign to become $\beta_1 = -2$.

6. The datafile MedGPA contains information on the medical school admission test (MCAT) scores for 55 students.

   (a) Fit a logistic regression model to predict Acceptance to medical school as a function of the number of schools a student applies to (Apps). Write down the prediction equations for the probability, odds, and log odds of acceptance using this model, filling in the numerical coefficients from the fitted model.

   (b) What would the fitted model say about the **probability** that a student who applies to four schools is accepted to medical school? What is this expressed in terms of **odds** of acceptance?

   (c) For approximately what number of applications would a student have roughly a 50/50 chance of being accepted to medical school, according to this model? (Hint: You might look at a graph, or solve one of the equations algebraically)