Math 150 - Methods in Biostatistics - Homework 4

your name here

Due: Wednesday, February 20, 2019, in class

```
knitr::opts_chunk$set(message=FALSE, warning=FALSE, fig.height=3, fig.width=5,
                      fig.align = "center")
library(tidyverse)
library(broom)
```

Note: there are two places to check for hints on R code. One is the class notes (http://st47s.com/Math150/Notes/, see R Examples) and the other is the R manual associated with the textbook which is on Sakai.

1. A1

Based on the description of the Challenger disaster O-ring concerns, identify which variable in the `Shuttle` data set in Table 7.1 should be the explanatory variable and which should be the response variable.

2. A2

Imagine you were an engineer working for Thiokol Corporation prior to January 1986. Create a few graphs of the data in Table 7.1. Is it obvious that temperature is related to the success of the O-rings? Submit any charts or graphs you have created that show a potential relationship between temperature and O-ring damage.

```
shuttle <- read_csv("~/Dropbox/teaching/math150/PracStatCD/Data Sets/Chapter 07/CSV Files/C7 Shuttle.csv
                    na="*")

# new names that make the data easier to work with:
# mine loads with an empty 5th column
names(shuttle) <- c("flight", "date", "temp", "launch", "X5")

# removing the row that has a missing value for launch
# also creating a character variable for success
shuttle <- shuttle %>%
  filter(!is.na(launch)) %>%
  mutate(launchsucc = ifelse(launch == 1, "success", "failure"))
```

3. A3

Use the data in Table 7.1 to create a scatterplot with a least squares regression line for the space shuttle data. Calculate the predicted response values ($\hat{y} = b_0 + b_1 x$) when the temperature is 60F and when the temperature is 85F.

4. A4

Solve Equation (7.5) for π_i to show that Equation (7.6) is true.

5. A5

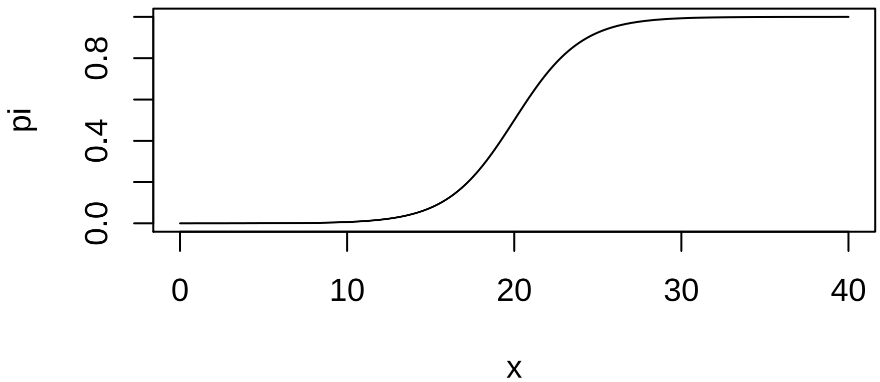
Use Equation (7.6) to create twelve graphs: In each graph plot the explanatory variable (x) versus the expected probability of success (pi) using the following values:

	1	2	3	4	5	6	7	8	9	10	11	12
β_0	-10	-10	-10	-5	-5	-5	10	10	10	5	5	5
β_1	0.5	1	1.5	0.5	1	1.5	-0.5	-1	-1.5	-0.5	-1	-1.5

(a) Do not submit the graphs, but explain the impact of changing β_0 and β_1.

(b) For all of the graphs, at what value of π does there appear to be the steepest slope?

```
exes <- seq(0,40,by=0.01)
b0 <- -10
b1 <- 0.5
plot(exes, exp(b0 + b1*exes) / (1 + exp(b0 + b1*exes)), type="l", xlab="x", ylab="pi")
```



6. A6

[For the shuttle data:] Use statistical software to calculate the maximum likelihood estimates of β_0 and β_1. Compare the maximum likelihood estimates to the least squares estimates in A3. Use `glm(response ~ explanatory, family = "binomial", data = yourdataset)`.

7. A7

Use Equation (7.9) to predict the probability that a launch has no O-ring damage when the temperature is 31F, 50F, and 75F.

8. A8

Calculate the odds of a launch with no O-ring damage when the temperature is 60F and when the temperature is 70F.

9. A9

When x_i increases by 10, state in terms of e^{b_1} how much you would expect the odds to change.

10. A10

The difference between the odds of success at 60F and 59F is about 0.3285 - 0.2605 = 0.068. Would you expect the difference between the odds at 52F and 51F to also be about 0.068? Explain why or why not.

11. A11

Create a plot of two logistic regression models. Plot temperature versus the estimated probability using maximum likelihood estimates from A6, and plot temperature versus the estimated probability using the least squares estimates from A3.