Homework 4A - Multiple Regression

STAT 242: Intermediate Statistics

The code below just loads some packages and makes it so that enough digits are printed that you won't get confused by rounding errors.

```
library(dplyr) # functions like summarize
library(ggplot2) # for making plots
library(readr)
library(Sleuth3)
options("pillar.sigfig" = 10) # print 10 significant digits in summarize output
```

Problem 1: Crabs (Adapted from Sleuth 3 exercise 9.17)

The description below comes from our book:

As part of a study of the effects of predatory intertidal crab species on snail populations, researchers measured the mean closing forces (in newtons) and the propodus heights (in mm) of the claws on several crabs of hreee species. (Data from S. B. Yamada and E. G. Boulding, "Claw Morphology, Prey Size Selection and Foraging Efficiency in Generalist and Specialist Shell-Breaking Crabs," *Journal of Experimental Marine Biology and Ecology, 220 (1998): 191-211.) Here we will examine the relationship between closing force (our response variable) and species and propodus height (explanatory variables). The following code reads the data in.

```
crabs <- ex0722
```

(a) Create an appropriate plot of the data involving all three variables. Does it appear that an additive model or a model with interactions between species and height would be more appropriate?

(b) Fit a multiple regression model to the data, allowing for different slopes for the different species. In this model, use the original Height and Force variables as explanatory and response variables, respectively. Create residual diagnostic plots of your model fit and calculate the standard deviation of the residuals within each group. Discuss any conditions for the regression model that are not satisfied.

(c) Find a set of transformations of the data so that the conditions of the multiple regression model are better satisfied (Note: I think you can do well enough with transformations of the response variable only). Verify that you have succeed by discussing residual diagnostic plots and standard deviations of the residuals across the different species. Recreate your plot of the data from part (a), but with your transformed variables this time.

(d) Write down the model you fit in part (c). This should not involve any numbers.

(e) Write down the equation for the estimated population mean (transformed) force as a function of species indicator variables and propodus height.

(f) Write down the equation for the estimated mean (transformed) forces as a function of propodus height, in the population of Lophopanopeus bellus crabs. Group together like terms so you have a single intercept and slope.

(g) What is the estimated change in (transformed) claw closing force that is associated with a 1 mm increase in propodus height, in the population of Cancer productus crabs? Just writing down a number is good enough.

(h) What is the estimated change in (transformed) claw closing force that is associated with a 1 mm increase in propodus height, in the population of Hemigrapsus Nudus crabs? Just writing down a number is good enough.

(i) Find and interpret a 95% confidence interval for the difference between the change in population mean (transformed) claw closing force that is associated with a 1 mm increase in propodus height in the populations of Hemigrapsus Nudus crabs and Cancer productus crabs. (That sentence was a lot to take in. I'm looking for a confidence interval for the difference between the population quantities from parts h and g.) Your answer should include a couple of sentences describing interpretation in context.

(j) Conduct a test of the claim that the slopes of lines describing the relationship between propodus height and (transformed) closing force is the same in the populations of crabs of all three species. State your null and alternative hypotheses in terms of model parameters, the p-value for the test, and your conclusion in context.