James Yang

DATA 558

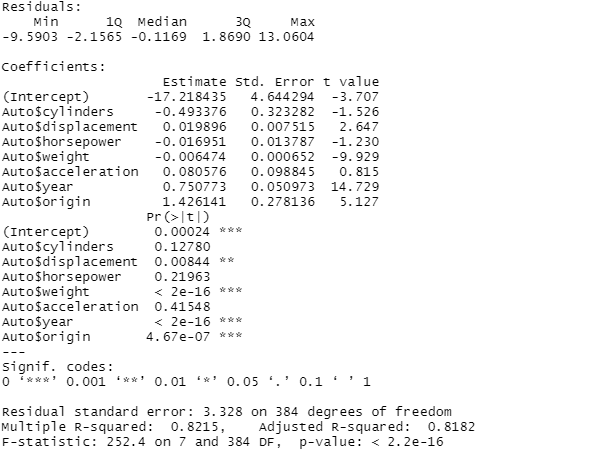
Homework 2

* 1. **Fit a least squares linear model to the data, in order to predict mpg using**

**all of the other predictors except for name. Present your results in the**

**form of a table. Be sure to indicate clearly how any qualitative variables**

**should be interpreted.**



When doing summary of the lm, the standard error is the variability to expect in the coefficients. So for example, for cylinders we can expect a variation of 0.323 cylinders. The T value is the coefficient/std.error. When looking at the Pr(>= |t|) values, we can see that if the values are greater than the standard alpha of 0.05, then we know that the coefficients are not statistically significant in our model. The residual standard error is 3.328, which means it can be off by an average of that amount when predicting mpg. Multiple R-Squared is the ratio of (1-(sum of squared error/sum of squared total)).

* 1. **What is the (training set) mean squared error of this model?**

Taking the residuals of the summary squared, we get: 10.84748

* 1. **What gas mileage do you predict for a Japanese car with three cylinders, displacement 100, horsepower of 85, weight of 3000, acceleration of 20, built in the year 1980?**

Text

Description automatically generated with low confidence

Multiplying these coefficients with the respective values above we get:

Mpg = 26.5

* 1. **On average, holding all other covariates fixed, what is the difference between the mpg of a Japanese car and the mpg of an American car?**

American cars get an average of 20.03 mpg and Japanese cars get an average of 30.45 mpg. The difference is 10.417 mpg.

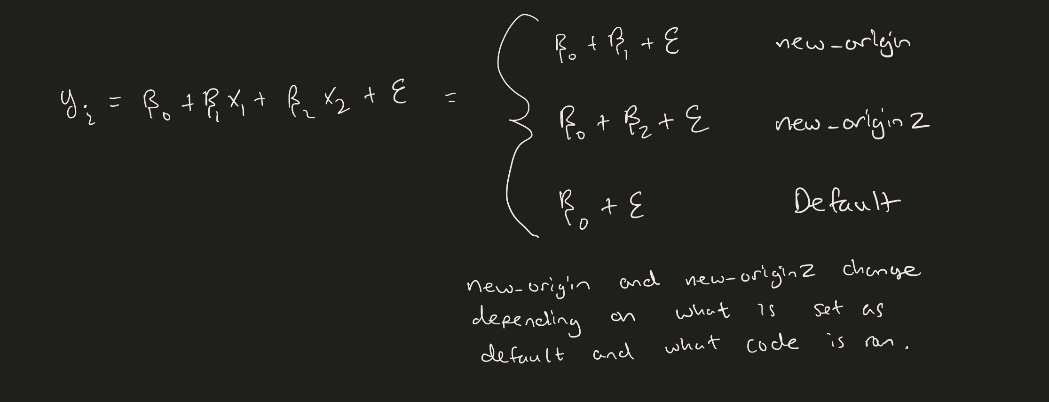
* 1. **On average, holding all other covariates fixed, what is the change in mpg associated with a 10-unit change in horsepower?**

\*\*\*\*\*\*INCOMPLETE\*\*\*\*\*\*

* 1. **First, code the origin variable using two dummy (indicator) variables,**

**with Japanese as the default value. Write out an equation like (3.30) in**

**the textbook, and report the coefficient estimates. What is the predicted**

**mpg for a Japanese car? for an American car? for a European car?**

If new origin german 1, else 0.

If new origin2 american 1, else 0.

A screenshot of a computer

Description automatically generated with medium confidence

American MPG: 20.033

Japanese MPG: 30.451

Germany MPG: 27.603

* 1. **Now, code the origin variable using two dummy (indicator) variables, with American as the default. Write out an equation like (3.30) in the textbook, and report the coefficient estimates. What is the predicted mpg for a Japanese car? for an American car? for a European car?**

**Text

Description automatically generated**

If new origin is Japanese 1, else 0.

If new origin 2 is german 1, else 0.

A screenshot of a computer

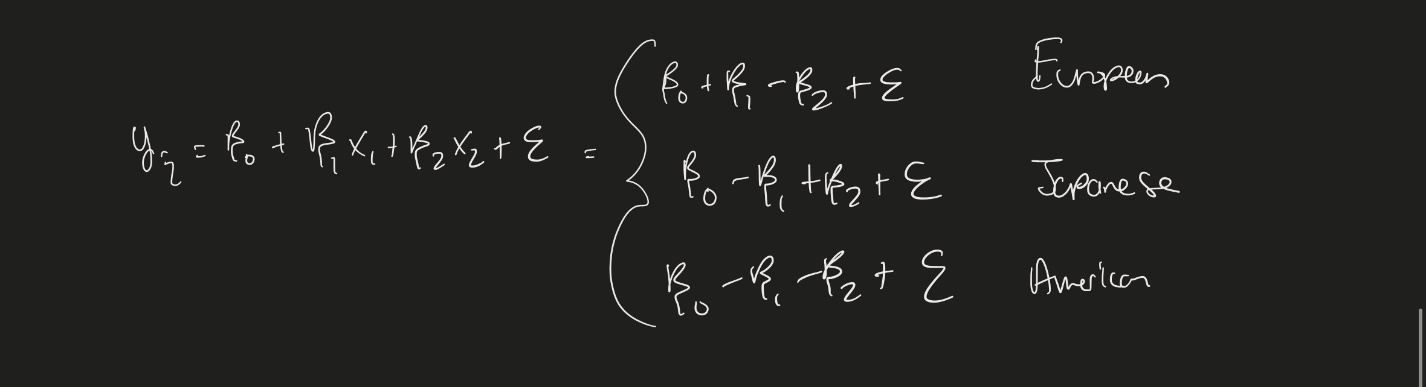
Description automatically generated with medium confidence

American MPG: 20.033

Japanese MPG: 30.451

Germany MPG: 27.603

* 1. **Now, code the origin variable using two variables that take on values of +1 or -1. Write out an equation like (3.30) in the textbook, and report the coefficient estimates. What is the predicted mpg for a Japanese car? for an American car? for a European car?**

****

If new origin is German, it will give 1 else -1.

If new origin2 is Japanese, it will give 1 else -1.

A screenshot of a computer

Description automatically generated with medium confidence

Japanese MPG: 27.6029

American MPG: 20.0335

Germany MPG: 29.0268

* 1. **Finally, code the origin variable using a single variable that takes on values of 0 for Japanese, 1 for American, and 2 for European. Write out an equation like (3.30) in the textbook, and report the coefficient estimates. What is the predicted mpg for a Japanese car? for an American car? For a European car?**

**Text

Description automatically generated**

Graphical user interface, text, application

Description automatically generated

A screenshot of a computer

Description automatically generated with low confidence

American MPG: 20.05

Japanese MPG: 25.0822

Germany MPG: 25.0822

* 1. **Comment on your results in (a)-(d).**

The values found in part c were a bit different than the ones found in the rest of the parts because there was a fluctuation in the values of points in which they were being put in. This caused a shift in values for the MPG. The cars that are American also were being valued less because they were no longer Japanese which meant that their MPG would get docked on the fact that they weren’t Japanese.

1. **Fit a model to predict mpg on the Auto dataset using origin and horsepower, as well as an interaction between origin and horsepower. Present your results, and write out an equation like (3.35) in the textbook. On average, how much does the mpg of a Japanese car change with a one-unit increase in horsepower? How about the mpg of an American car? a European car?**

.

1. For these questions, I have used the gcookbook dataset with a random set of height and weight in it.
   1. When solving the general equation, we get **Y = 142.1.**
   2. When plugging in a gcookbook heightweight dataset and performing an lm on it, I get that   
      Y = 44.867x1 – 128.328. For someone to be 5.33 feet, we predict **Y = 110.96.**
   3. Y =

**=**

We know the value of the beta coefficient is 4.8.

* 1. According to the test set from gcookbook, the MSE values for all 3 models appear to be the same. The reason is because the error found is just going to be a factor of 12, meaning their results should still differentiate the same and average out to be the same as well.

Text

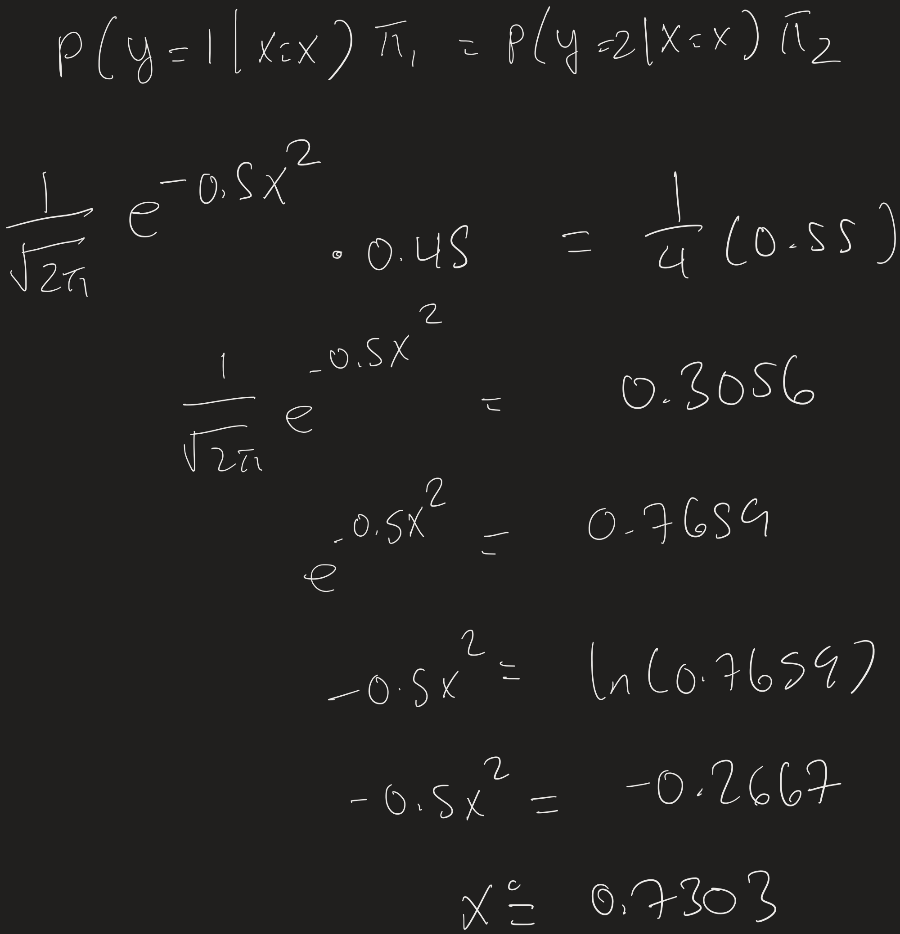
Description automatically generated



A picture containing arrow

Description automatically generated

A picture containing text

Description automatically generated

**+- 0.703**

A picture containing chart

Description automatically generated

**Figure shows the intersection and the two distributions.**

* 1. We can use the training observations to estimate n sample mean, sample SD, and pi values relating to class 1.
  2. Not done yet