STA 4320 HW 3

Mary Robinson

9/25/24

Question 1

Part a

Loading data and finding the first six rows

```
# install packages
require(ISLR2)
```

```
## Loading required package: ISLR2
```

```
dat = Auto[1:5]
head(dat)
```

```
##   mpg cylinders displacement horsepower weight
## 1  18         8          307        130   3504
## 2  15         8          350        165   3693
## 3  18         8          318        150   3436
## 4  16         8          304        150   3433
## 5  17         8          302        140   3449
## 6  15         8          429        198   4341
```

Part b

Regression Output

```
# reg = lm(mpg ~ cylinders + displacement + horsepower + weight, data = dat)
#or
reg = lm(mpg ~ ., data = dat)
summary(reg)
```

```
##
## Call:
## lm(formula = mpg ~ ., data = dat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -11.5248  -2.7964  -0.3568   2.2577  16.3221
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)  45.7567705  1.5200437  30.102  < 2e-16 ***
## cylinders    -0.3932854  0.4095522  -0.960 0.337513
## displacement  0.0001389  0.0090099   0.015 0.987709
## horsepower   -0.0428125  0.0128699  -3.327 0.000963 ***
## weight       -0.0052772  0.0007166  -7.364 1.08e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.242 on 387 degrees of freedom
## Multiple R-squared:  0.7077, Adjusted R-squared:  0.7046
## F-statistic: 234.2 on 4 and 387 DF,  p-value: < 2.2e-16
```

```
#percentage of variation in Y is Multiple R-squared value
```

Part c

Prediction Interval #for this one we are looking for a SINGLE point (car), so prediction interval instead of confidence interval

```
pred_int = predict(reg, data.frame(cylinders = 6, displacement = 275, horsepower = 12
5, weight = 3500, interval = "prediction", level = 0.95))
pred_int_rounded = round(as.numeric(pred_int[2:3]))

paste("The prediction interval is ", paste(pred_int_rounded, collapse = " "))
```

```
## [1] "The prediction interval is  NA NA"
```

Part d

H0: beta_1 = beta_2 = beta_3 = beta_4 = 0

H1: beta_j is nonzero for at least one of j= 1,2,3,4

The p value to the F Test less than 2.2e^-16. So we reject H0 at the 0.05 level.

Part e

with the intercept, cylinder, horsepower, and weight already in the model

H0: beta_2 = 0

H1: beta_2 is not equal to 0

(beta_2 is the slope for displacement)

The p value for the T test is 0.9877, way larger than 0.05, so we fail to reject H0.

Thus with the intercept, cylinder, horsepower, and weight already in the model, there is no significant evidence that we should also include displacement.

Once completed, you can compile into pdf if you have LaTeX, or compile to html and then print that html file into a pdf file if you do not have LaTeX.

Part f

```
round(cor(dat), 4)
```

```
##                  mpg cylinders displacement horsepower  weight
## mpg           1.0000   -0.7776      -0.8051    -0.7784 -0.8322
## cylinders    -0.7776    1.0000       0.9508     0.8430  0.8975
## displacement -0.8051    0.9508       1.0000     0.8973  0.9330
## horsepower   -0.7784    0.8430       0.8973     1.0000  0.8645
## weight       -0.8322    0.8975       0.9330     0.8645  1.0000
```

Part g

full model

```
y= Auto$mpg
n=392
```

RSE= sqrt(RSS / (n-p-1)) RSE: residual standard error

4.242 = sqrt(RSS / (392-4-1)) RSS= 6963.896

Sub model (with onlu horsepower and weight and intercept)

```
reg_0 = lm(mpg ~ horsepower + weight , data= dat)
RSS_0 = sum( ( y-predict(reg_0, data.frame(dat[,c(4,5)])) )^2)
```

compute the test statistic

F_0 = ((RSS_0 - RSS)/ q) / (RSS/ (n-p-1)) F_0 - (6993.845- 6963.896) / 2) / (6963.896 / (392-4-1)) F_0 = 0.8322

Compute the p value

```
p_val = 1- pf(0.8322, 2, 387 )
p_val
```

```
## [1] 0.4358681
```

```
# degrees of freedoms: q=2, n-p-1=392-4-1= 387
```

Make a conclusion we fail to reject H0 at the 0.01 level as the pvalue is way larger than 0.01

Therefore, with the intercept, horsepower, and wreight already in the model, there is insufficient ecidence that we should include cylinders and displacement.