HW2 Cranmer and Liu

Evan Cranmer and Xiaoqing Liu

2024-09-13

Q1 Creating a regression model

```
library(ISLR2)
?lm
mod1 <- lm(mpg ~ acceleration, data = Auto)

#print the model summary
summary(mod1)
```

```
##
## Call:
## lm(formula = mpg ~ acceleration, data = Auto)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -17.989  -5.616  -1.199   4.801  23.239
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    4.8332     2.0485   2.359   0.0188 *
## acceleration   1.1976     0.1298   9.228   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.08 on 390 degrees of freedom
## Multiple R-squared:  0.1792, Adjusted R-squared:  0.1771
## F-statistic: 85.15 on 1 and 390 DF,  p-value: < 2.2e-16
```

Q2 and Q3 How would you interpret the intercept and slope from 1

The model predicts that when acceleration is 0, 4.8332 miles per gallon is expected Every one unit increase in acceleration corresponds to a 1.1976 increase in miles per gallon. The intercept does not make sense to me because if a car is not moving it should not have that poor of miles per gallon.

Refitting the model

```
#centering around acceleration
library(ISLR2)
Auto$acc_cen <- Auto$acceleration - mean(Auto$acceleration)

summary(lm(mpg ~ acc_cen, Auto))
```

```
##
## Call:
## lm(formula = mpg ~ acc_cen, data = Auto)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -17.989  -5.616  -1.199   4.801  23.239
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  23.4459     0.3576  65.564   <2e-16 ***
## acc_cen       1.1976     0.1298   9.228   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.08 on 390 degrees of freedom
## Multiple R-squared:  0.1792, Adjusted R-squared:  0.1771
## F-statistic: 85.15 on 1 and 390 DF,  p-value: < 2.2e-16
```

The new intercept is saying that the expected mpg for cars with a mean level of acceleration is 23.45.

Question 4

The null hypothesis states that acceleration has no effect on mpg. This would not change, even if we centered the model, because the only thing we are doing is shifting the intercept to make it easier to interpret. Our null stays the same.

Question 5

A more flexible modeling approach can make it easier to interpret the intercept. If we were showing a presentation to clients, it might be wise to make a model that is easier to interpret for them. A trade off of centering the model is that the regression line can fit differently around the data points, even if we have the same slope.