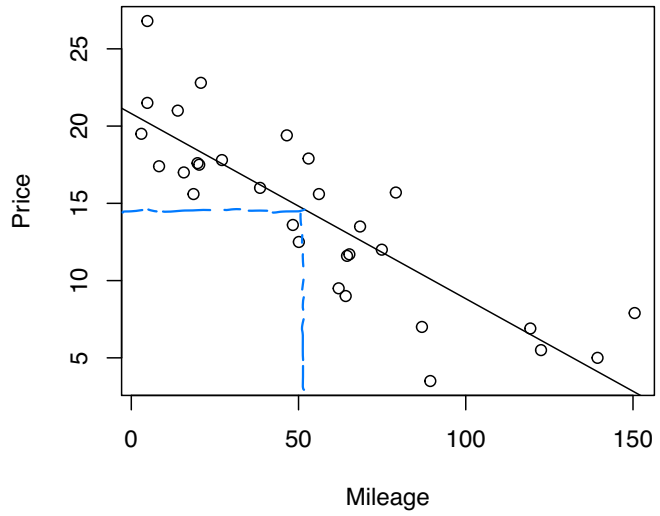
Lecture 12: Simple Linear Regression
Practice Problems
STAT 310, Spring 2021

The following scatterplot shows the association between price (in $1,000's) and mileage (number of miles driven in 1,000's) for a sample of 30 used Honda Accords in 2017. Also provided below is the output from fitting a simple linear regression model in R.



$\hat{\beta}_0 = 20.81$

$\hat{\beta}_1 = -0.12$

```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  20.8096     0.9529   21.84  < 2e-16 ***
Mileage      -0.1198     0.0141   -8.50 3.06e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.085 on 28 degrees of freedom
Multiple R-squared:  0.7207,Adjusted R-squared:  0.7107
F-statistic: 72.25 on 1 and 28 DF,  p-value: 3.055e-09
```

(a) Describe the association between price and mileage.

Negative linear association

(b) What are the explanatory and response variables for the linear regression model?

Explanatory variable: Mileage (x)

Response variable: Price (y)

(c) Write the equation for the least squares line.

$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x = 20.81 - 0.12x$

(d) What is the predicted price for a used Honda Accord that has been driven 50 thousand miles?

$\hat{y} = 20.81 - 0.12(50) = 14.81$

The prediction is \$14,810

(e) Interpret the slope.

An increase in # of miles driven by 1000 is associated with a decrease in price by $0.12 * 1000 = 120$ dollars.

(f) Interpret the intercept.

A Honda Accord that has been driven 0 miles, is predicted to cost 20,810 dollars.

(g) Calculate the residual for a car, in this data set, that costs 3.5 thousand dollars, and has been driven 89.4 thousand miles.

$e_i = y_i - \hat{y}_i = 3.5 - [20.81 - 0.12(89.4)]$

$= 3.5 - 10.08 = \boxed{-6.58}$

(h) Interpret the R^2. $R^2 = 0.72$

About 72% of the variability in car price can be explained by # of miles driven.