Day7

Olivia Wu

2024-03-04

Problem 3.1

a) A student who got perfect on the midterm and project would have a predicted score of 100 on the final.

b)

$$\hat{Final} = 11.0 + 0.53(87) + 1.20(21) = 82.31$$

$$Residual = 80 - 82.31 = \boxed{-2.31}$$

Michael's final grade was 2.31 points less than what was predicted by the model.

Problem 3.3

No, the equation only suggests that a point increase in the midterm grade tends to reflect in 0.53 points in the final grade, assuming the project grade is held constant. It says nothing about the strength of the relationship.

Problem 3.5

A one point increase in the project grade is associated with a 1.2 point increase in the final grade, given that the midterm grade does not change.

Problem 3.7

a) True

b) False

Problem 3.17

a) $H_0 : \beta_2 = 0$

$H_a : \beta_2 \neq 0$

It is given that all conditions are met.

The test statistic given by output is $t = 1.08$. With 228 degrees of freedom, $p = 2 \cdot P(t > 1.08) = 2 \cdot 0.1406 = 0.282$. Since $p > 0.05$, we fail to reject the null hypothesis. There is not enough evidence to suggest a linear relationship between weight and active pulse rates.

b) We are 90% confident that as the weight increases by 1 pound, the increase in active pulse rate lies between -0.0182 and 0.0866, assuming all other variables stay the same.

```
## [1] "(-0.0182041329565737, 0.0866041329565737)"
```

c)

$$\hat{Active} = 11.8 + 1.12(76) + 0.0342(200) - 1.09(7) = \boxed{96.13 \text{ bpm}}$$

Problem 3.18

a)

```
##
## Call:
## lm(formula = adj2007 ~ distance, data = RT)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -190.55  -58.19  -17.48   25.22  444.41
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  388.204     14.052  27.626  < 2e-16 ***
## distance     -54.427      9.659  -5.635 1.56e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 92.13 on 102 degrees of freedom
## Multiple R-squared:  0.2374, Adjusted R-squared:  0.2299
## F-statistic: 31.75 on 1 and 102 DF,  p-value: 1.562e-07
```

b) The estimated coefficients of *distance* are -54.427 and -16.486, and the R^2 values are 0.2374 and 0.7655. The second model really improved the R^2 value.

```
##
## Call:
## lm(formula = adj2007 ~ distance + squarefeet, data = RT)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -138.835  -32.621   -1.903   27.369  145.504
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  109.742     20.057   5.472 3.25e-07 ***
## distance     -16.486      5.942  -2.775  0.00659 **
## squarefeet   150.780      9.998  15.080  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 51.34 on 101 degrees of freedom
## Multiple R-squared:  0.7655, Adjusted R-squared:  0.7608
## F-statistic: 164.8 on 2 and 101 DF,  p-value: < 2.2e-16
```

c) In the simple regression model, we expect for every mile increase in distance to a trail, the house price increases by a value between -\$73.41 and -\$35.45. If we adjust for the house size, we expect the increase to be between -\$28.27 and -\$4.70. By adding $squarefeet$ into the regression, the width of the confidence interval decreased and the values shifted higher.

```
## [1] "(-73.4070661291882, -35.4469338708118)"
```

```
## [1] "(-28.2733296195622, -4.69867038043779)"
```

d)

$$\hat{Price} = 109.742 - 16.486(0.5) + 150.78(1500) = 226,271.499$$

We expect the house to cost \$266,271.