Stat 21 Homework 6

Solution for Problem 2

Problem 2

a) Build a linear regression model relating gasoline mileage, y, to vehicle weight x_3 and the type of transmission x_2 . Does the type of transmission significantly affect the mileage performance? Justify your answer.

b) Modify the model developed in part c to include an interaction between vehicle weight and the type of transmission. What is the average effect on gasoline mileage when the transmission is automatic? What is the average effect on gasoline mileage when the transmission is manual?

Solution for a)

```
MLR_car2 <- lm(mpg ~ weight + transmission_type,
               data=mileage)
summary(MLR_car2)
```

```
##
## Call:
## lm(formula = mpg ~ weight + transmission_type, data = mileage)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.2095 -2.2586  0.3033  2.2403  7.0699
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)       34.2056710  3.8447835   8.897 8.73e-10 ***
## weight            -0.0042267  0.0009466  -4.465 0.000112 ***
## transmission_typeM  3.7157618  1.9791784   1.877 0.070552 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.354 on 29 degrees of freedom
## Multiple R-squared:  0.7364, Adjusted R-squared:  0.7182
## F-statistic:  40.5 on 2 and 29 DF,  p-value: 4.025e-09
```

Note we have fit the following regression model:

$$\hat{y} = \hat{\beta}_0 + (\hat{\beta}_1 \times weight) + (\hat{\beta}_2 \times type)$$

where $type$ is a dummy variable with $type = 0$ indicating an automatic and $type = 1$ indicating a manual.

Let $\alpha = 0.05$. Based on the individual t-test for the signifigance of the difference in average mileage between manual and automatic cars in this model, no transmission does not have a significant effect on mileage, given that weight is included in the model, since p-value$= 0.071 > \alpha$.

Solution for b)

```r
MLR_car2_interaction <- lm(mpg ~ weight + transmission_type +
                             weight*transmission_type, data=mileage)
summary(MLR_car2_interaction)
```

```
##
## Call:
## lm(formula = mpg ~ weight + transmission_type + weight * transmission_type,
##     data = mileage)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.4534 -1.8453  0.3717  1.4173  4.9229
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)              29.4530696  3.2887177   8.956 1.04e-09 ***
## weight                   -0.0030367  0.0008114  -3.743 0.000834 ***
## transmission_typeM       28.6553504  6.2299643   4.600 8.28e-05 ***
## weight:transmission_typeM -0.0094807  0.0022902  -4.140 0.000289 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.689 on 28 degrees of freedom
## Multiple R-squared:  0.8365, Adjusted R-squared:  0.8189
## F-statistic: 47.73 on 3 and 28 DF,  p-value: 3.908e-11
```

Note we have fit the following regression model:

$$\hat{y} = \hat{\beta}_0 + (\hat{\beta}_1 \times weight) + (\hat{\beta}_2 \times type) + (\hat{\beta}_3 \times weight \times type)$$

where $type$ is a dummy variable with $type = 0$ indicating an automatic and $type = 1$ indicating a manual.

There is a significant interaction between car weight and the type of transmission. For an automatic transmission type, the predicted mileage per gallon is $29.4530696 - (0.0030367 \times weight)$ which indicates that, *on average*, for every pound increase in weight miles per gallon decreases on average by 0.003. For an manual transmission type, the predicted mileage per gallon is $(29.4530696 + 28.6553504) + ((-0.0030367 - 0.0094807) \times weight) = 58.10842 - (0.0125174 \times weight)$ which indicates that, *on average*, for every pound increase in weight, miles per gallon decreases on average by 0.013.

Standardized mileage data solutions

a)

```r
MLR_car3 <- lm(scale(mpg) ~ weight + transmission_type,
               data=mileage)
summary(MLR_car3)
```

```
##
## Call:
## lm(formula = scale(mpg) ~ weight + transmission_type, data = mileage)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.9828 -0.3575  0.0480  0.3546  1.1190
##
## Coefficients:
```

```
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)          2.2130273  0.6085166   3.637 0.001062 **
## weight              -0.0006690  0.0001498  -4.465 0.000112 ***
## transmission_typeM   0.5880962  0.3132459   1.877 0.070552 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5309 on 29 degrees of freedom
## Multiple R-squared:  0.7364, Adjusted R-squared:  0.7182
## F-statistic:  40.5 on 2 and 29 DF,  p-value: 4.025e-09
```

Note we have fit the following regression model:

$$\hat{y} = \hat{\beta}_0 + (\hat{\beta}_1 \times weight) + (\hat{\beta}_2 \times type)$$

where *type* is a dummy variable with $type = 0$ indicating an automatic and $type = 1$ indicating a manual.

Let $\alpha = 0.05$. Based on the individual t-test for the signifigance of the difference in average mileage between manual and automatic cars in this model, no transmission does not have a significant effect on mileage, given that weight is included in the model, since p-value$= 0.071 > \alpha$.

b)

```r
MLR_car3_interaction <- lm(scale(mpg) ~ weight + transmission_type +
                             weight*transmission_type, data=mileage)
summary(MLR_car3_interaction)
```

```
##
## Call:
## lm(formula = scale(mpg) ~ weight + transmission_type + weight *
##     transmission_type, data = mileage)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.86311 -0.29206  0.05883  0.22432  0.77915
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 1.4608297  0.5205076   2.807 0.009012 **
## weight                     -0.0004806  0.0001284  -3.743 0.000834 ***
## transmission_typeM          4.5353022  0.9860208   4.600 8.28e-05 ***
## weight:transmission_typeM -0.0015005  0.0003625  -4.140 0.000289 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4255 on 28 degrees of freedom
## Multiple R-squared:  0.8365, Adjusted R-squared:  0.8189
## F-statistic: 47.73 on 3 and 28 DF,  p-value: 3.908e-11
```

Note we have fit the following regression model:

$$\hat{y} = \hat{\beta}_0 + (\hat{\beta}_1 \times weight) + (\hat{\beta}_2 \times type) + (\hat{\beta}_3 \times weight \times type)$$

where *type* is a dummy variable with $type = 0$ indicating an automatic and $type = 1$ indicating a manual.

There is a significant interaction between vehicle weight and the type of transmission. For an automatic transmission type, the predicted (standardized) mileage is $1.461 - (0.0004806 \times weight)$ which indicates that, *on average*, for every pound increase in weight, standardized mileage decreases by 0.0005. For an manual transmission type, the predicted mileage per gallon is $(1.4608297 + 4.5353022) + ((-0.0004806 - 0.0015005) \times \mathit{weight}) = 5.996132 - (0.0019811 \times \mathit{weight})$ which indicates that, *on average*, for every pound increase in weight, standardized mileage decreases by 0.002.