HW1-Q5

Zhaoqi Li

2023/9/30

```
library(ISLR)
```

a.

```
model1 <- lm(Sales ~ Price + Urban + US, data = Carseats)
```

b.

```
summary(model1)
```

```
##
## Call:
## lm(formula = Sales ~ Price + Urban + US, data = Carseats)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.9206 -1.6220 -0.0564  1.5786  7.0581
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 13.043469   0.651012  20.036  < 2e-16 ***
## Price       -0.054459   0.005242 -10.389  < 2e-16 ***
## UrbanYes    -0.021916   0.271650  -0.081    0.936
## USYes        1.200573   0.259042   4.635 4.86e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.472 on 396 degrees of freedom
## Multiple R-squared:  0.2393, Adjusted R-squared:  0.2335
## F-statistic: 41.52 on 3 and 396 DF,  p-value: < 2.2e-16
```

Interpretation

Price: If other variables are fixed (Urban and USYes), when the price of carseats increase 1 dollar, the sales of car seats will decrease 54.46 on average.

UrbanYes: If other variables are fixed (Price and USYes), when the location is in urban, the sales of car seats will decrease 21.91 on average than that is not in urban.

USYes: If other variables are fexed (Price and UrbanYes), when the location is in US, the sales of car seats will increase 1201 on average than that is not in US.

c.

$$Sales_i = 13.04 - 0.054 \times Price_i - 0.021 \times Urban_i + 1.2 \times US_i$$

$Urban_i$: If the location is in urban i=1; if the location is not in urban i=0.

US_i:If the location is in US i=1; if the location is not in US i=0.

d.

According to the output, the p value of Price and USYes are significant at 5% level, for these two predictor we can reject H0.

e.

```
model2 <- lm(Sales ~ Price + US, data = Carseats)
```

f.

```
summary(model1)$r.squared
```

```
## [1] 0.2392754
```

```
summary(model2)$r.squared
```

```
## [1] 0.2392629
```

The R^2 of these two models are similar and that of model 1 is a little large. No, we should also consider adjusted R-square which consider the complexity of the model. For the R-square, it determines the proportion of variance in the dependent variable that can be explained by the independent variable. Therefore, the increase of covariates will have larger R-square ### g.

```
confint(model2)
```

```
##                   2.5 %      97.5 %
## (Intercept) 11.79032020 14.27126531
## Price       -0.06475984 -0.04419543
## USYes        0.69151957  1.70776632
```

h.

```
model3 <- lm(data=Carseats, Sales ~ Price*US)
summary(model3)
```

```
##
## Call:
## lm(formula = Sales ~ Price * US, data = Carseats)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.9299 -1.6375 -0.0492  1.5765  7.0430
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 12.974798   0.953079  13.614  < 2e-16 ***
## Price       -0.053986   0.008163  -6.613 1.22e-10 ***
## USYes        1.295775   1.252146   1.035    0.301
## Price:USYes -0.000835   0.010641  -0.078    0.937
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.472 on 396 degrees of freedom
## Multiple R-squared:  0.2393, Adjusted R-squared:  0.2335
## F-statistic: 41.52 on 3 and 396 DF,  p-value: < 2.2e-16
```

Interpretation

Price: When the price of carseats increase 1 dollar, the sales of car seats will decrease 53.99 on average for a non-US store.

USYes: When the price is zeor, the sales of car seats for US store is 1295 more on average than that for a non-US store.

Price:USYes: When the price of carseats increase 1 dollar for US stores in US,the sales of car seat will 0.84 fewer than than that for a non-US store.