# ST_503 HW2

Halid Kopanski

2022-05-30

**LMR 3.4 a)**

```
##
## Call:
## lm(formula = total ~ expend + ratio + salary, data = sat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -140.911  -46.740   -7.535   47.966  123.329
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1069.234    110.925   9.639 1.29e-12 ***
## expend        16.469     22.050   0.747   0.4589
## ratio          6.330      6.542   0.968   0.3383
## salary        -8.823      4.697  -1.878   0.0667 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 68.65 on 46 degrees of freedom
## Multiple R-squared:  0.2096, Adjusted R-squared:  0.1581
## F-statistic: 4.066 on 3 and 46 DF,  p-value: 0.01209
```

| Res.Df | RSS | Df | Sum of Sq | F | Pr(>F) |
|---|---|---|---|---|---|
| 49 | 274307.7 | NA | NA | NA | NA |
| 46 | 216811.9 | 3 | 57495.74 | 4.066203 | 0.0120861 |

| | Df | Sum Sq | Mean Sq | F value | Pr(>F) |
|---|---|---|---|---|---|
| expend | 1 | 39722.059 | 39722.059 | 8.4276482 | 0.0056579 |
| ratio | 1 | 1142.671 | 1142.671 | 0.2424354 | 0.6247946 |
| salary | 1 | 16631.014 | 16631.014 | 3.5285264 | 0.0666677 |
| Residuals | 46 | 216811.935 | 4713.303 | NA | NA |

| df_vector | SS | MS | F_stat | p_value |
|---|---|---|---|---|
| 3 | 57495.74 | 19165.248 | 4.066203 | 0.0120861 |
| 46 | 216811.94 | 4713.303 | NA | NA |
| 49 | 274307.68 | NA | NA | NA |

There is not enough evidence to reject the null hypothesis that $\beta_{salary}$ is zero. The p value is above the critical value of 0.025 for a two sided t test.

There is sufficient evidence to reject the null hypothesis that states $\beta_{expend} = \beta_{ratio} = \beta_{salary}$. The f statistic and the corresponding p value are 4.0662033 and 0.0120861 respectively.

Given the large p values and low $R^2$, we cannot say that the predictors had a large effect on the response variable.

**LMR 3.4 b)**

```
##
## Call:
## lm(formula = total ~ expend + ratio + salary + takers, data = sat)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -90.531 -20.855  -1.746  15.979  66.571
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1045.9715    52.8698  19.784  < 2e-16 ***
## expend         4.4626    10.5465   0.423    0.674
## ratio         -3.6242     3.2154  -1.127    0.266
## salary         1.6379     2.3872   0.686    0.496
## takers        -2.9045     0.2313 -12.559 2.61e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 32.7 on 45 degrees of freedom
## Multiple R-squared:  0.8246, Adjusted R-squared:  0.809
## F-statistic: 52.88 on 4 and 45 DF,  p-value: < 2.2e-16
```

| Res.Df | RSS | Df | Sum of Sq | F | Pr(>F) |
|---|---|---|---|---|---|
| 49 | 274307.7 | NA | NA | NA | NA |
| 45 | 48123.9 | 4 | 226183.8 | 52.87534 | 0 |

| Res.Df | RSS | Df | Sum of Sq | F | Pr(>F) |
|---|---|---|---|---|---|
| 46 | 216811.9 | NA | NA | NA | NA |
| 45 | 48123.9 | 1 | 168688 | 157.7379 | 0 |

| | Df | Sum Sq | Mean Sq | F value | Pr(>F) |
|---|---|---|---|---|---|
| expend | 1 | 39722.059 | 39722.059 | 37.143557 | 0.0000002 |
| ratio | 1 | 1142.671 | 1142.671 | 1.068496 | 0.3068088 |
| salary | 1 | 16631.014 | 16631.014 | 15.551435 | 0.0002779 |
| takers | 1 | 168688.039 | 168688.039 | 157.737888 | 0.0000000 |
| Residuals | 45 | 48123.896 | 1069.420 | NA | NA |

| df_vector | SS | MS | F_stat | p_value |
|---|---|---|---|---|
| 4 | 226183.8 | 56545.95 | 52.87534 | 0 |
| 45 | 48123.9 | 1069.42 | NA | NA |
| 49 | 274307.7 | NA | NA | NA |

We have sufficient evidence to reject the null hypothesis that $\beta_{takers} = 0$ at the $\alpha$ level of 0.05 (p value in this case is close to zero).

**LMR 3.7 a)**

```
##
## Call:
## lm(formula = Distance ~ RStr + LStr + RFlex + LFlex, data = punting)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -23.941  -8.958  -4.441  13.523  17.016
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -79.6236    65.5935  -1.214    0.259
## RStr          0.5116     0.4856   1.054    0.323
## LStr         -0.1862     0.5130  -0.363    0.726
## RFlex         2.3745     1.4374   1.652    0.137
## LFlex        -0.5277     0.8255  -0.639    0.541
##
## Residual standard error: 16.33 on 8 degrees of freedom
## Multiple R-squared:  0.7365, Adjusted R-squared:  0.6047
## F-statistic:  5.59 on 4 and 8 DF,  p-value: 0.01902
```

None of the predictors are significant at the 5% level

**LMR 3.7 b)**

```
##
## Call:
## lm(formula = Distance ~ 1, data = punting)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -43.303  -7.983   1.937  15.267  43.767
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  148.233      7.203   20.58 9.99e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 25.97 on 12 degrees of freedom
```

```
## Analysis of Variance Table
##
## Model 1: Distance ~ 1
## Model 2: Distance ~ RStr + LStr + RFlex + LFlex
##   Res.Df    RSS Df Sum of Sq      F  Pr(>F)
## 1     12 8093.3
## 2      8 2132.6  4    5960.7 5.5899 0.01902 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

| df_vector | SS | MS | F_stat | p_value |
|---|---|---|---|---|
| 4 | 5960.668 | 1490.1669 | 5.589941 | 0.0190248 |
| 8 | 2132.641 | 266.5801 | NA | NA |
| 12 | 8093.308 | NA | NA | NA |

**LMR 3.7 c)**

```
##
## Call:
## lm(formula = Distance ~ RStr + LStr, data = punting)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -29.280  -9.583   3.147  10.266  26.450
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  12.8490    33.0334   0.389    0.705
## RStr          0.7208     0.4913   1.467    0.173
## LStr          0.2011     0.4883   0.412    0.689
##
## Residual standard error: 17.24 on 10 degrees of freedom
## Multiple R-squared:  0.6327, Adjusted R-squared:  0.5592
## F-statistic: 8.611 on 2 and 10 DF,  p-value: 0.00669
```

**LMR 3.7 d)**

```
##                    2.5 %     97.5 %
## (Intercept) -230.8826019 71.6353107
## RStr          -0.6080871  1.6313618
## LStr          -1.3690973  0.9966981
## RFlex         -0.9400454  5.6890474
## LFlex         -2.4313762  1.3759086
```