Linear Model

Calista Densing
December 6, 2019

```
myData = read.csv('rawData.csv')
```

H_0: No difference in the models
H_A : There is a difference in the models

First model:

SIMS is a function of ARM

```
model1 = lm(SIMS~ARM, data=myData)
summary(model1)
```

```
##
## Call:
## lm(formula = SIMS ~ ARM, data = myData)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.6467 -0.7502 -0.0285  0.6875  3.0770
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -4.095160   0.391745  -10.45   <2e-16 ***
## ARM          0.054563   0.004806   11.35   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.226 on 145 degrees of freedom
## Multiple R-squared:  0.4706, Adjusted R-squared:  0.467
## F-statistic: 128.9 on 1 and 145 DF,  p-value: < 2.2e-16
```

predict SIMS for ARM = 88

```
x = data.frame(GRIP=94, ARM=88)
predSIMS1 = predict.lm(model1,x)
print(predSIMS1)
```

```
##         1
## 0.7063836
```

prediction interval:

```
predict(model1, x,interval= 'predict')
```

```
##         fit       lwr      upr
## 1 0.7063836 -1.726209 3.138977
```

Second model:

SIMS is a function of GRIP

```
model2 = lm(SIMS~GRIP, data=myData)
summary(model2)
```

```
##
## Call:
## lm(formula = SIMS ~ GRIP, data = myData)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.9295 -0.8708 -0.1219  0.8039  3.3494
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -4.809675   0.511141   -9.41   <2e-16 ***
## GRIP         0.045463   0.004535   10.03   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.295 on 145 degrees of freedom
## Multiple R-squared:  0.4094, Adjusted R-squared:  0.4053
## F-statistic: 100.5 on 1 and 145 DF,  p-value: < 2.2e-16
```

predict SIMS for GRIP = 94

```
predSIMS2 = predict.lm(model2,x)
print(predSIMS2)
```

```
##          1
## -0.5361543
```

prediction interval:

```
predict(model2, x,interval= 'predict')
```

```
##          fit       lwr      upr
## 1 -0.5361543 -3.107961 2.035652
```

Third model:

SIMS is a prediction of ARMS and GRIP

```
model3 = lm(SIMS~ARM+GRIP, data=myData)
summary(model3)
```

```
##
## Call:
## lm(formula = SIMS ~ ARM + GRIP, data = myData)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.1846 -0.7380 -0.0641  0.7394  2.8634
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -5.433871   0.461815 -11.766  < 2e-16 ***
## ARM          0.037311   0.005774   6.462 1.50e-09 ***
## GRIP         0.024470   0.005158   4.744 4.99e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.144 on 144 degrees of freedom
## Multiple R-squared:  0.5422, Adjusted R-squared:  0.5358
## F-statistic: 85.26 on 2 and 144 DF,  p-value: < 2.2e-16
```

predict SIMS for ARM=88 and GRIP=94

```
predSIMS3 = predict.lm(model3,x)
print(predSIMS3)
```

```
##         1
## 0.1496476
```

prediction interval:

```
predict(model3, x,interval= 'predict')
```

```
##         fit       lwr      upr
## 1 0.1496476 -2.132373 2.431668
```

comparison between models 1 and 3

```
anova(model1,model3)
```

```
## Analysis of Variance Table
##
## Model 1: SIMS ~ ARM
## Model 2: SIMS ~ ARM + GRIP
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1    145 217.88
## 2    144 188.43  1     29.45 22.506 4.994e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Conclusion:

Since the p-value = 4.994e-06, we reject the null hypothesis. There is a difference between model 1 https://github.com/calistaelyse/LinearModelWriteup.gitnd model 3 because the p-value is less than 0.05. The residual sums of model squares from model 1 is less than the residual model squares from model 3. This means that model 3 is a better fit.