Assignment 3

Shinjon Ghosh

2025-01-31

```r
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
data(mtcars)
```

```r
1.## Full Model Summary
full_model <- lm(mpg ~ cyl + hp + drat + wt, data = mtcars)
summary(full_model)
```

```
##
## Call:
## lm(formula = mpg ~ cyl + hp + drat + wt, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.6171 -1.5663 -0.6058  1.2612  5.8161
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 34.49588    7.44101   4.636  8.1e-05 ***
## cyl         -0.76229    0.63502  -1.200  0.24040
## hp          -0.02089    0.01295  -1.613  0.11845
## drat         0.81771    1.38684   0.590  0.56034
## wt          -2.97331    0.81818  -3.634  0.00116 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.541 on 27 degrees of freedom
## Multiple R-squared:  0.8451, Adjusted R-squared:  0.8222
## F-statistic: 36.84 on 4 and 27 DF,  p-value: 1.438e-10
```

```r
## SSE, SSR, SSTO
SSE <- sum(residuals(full_model)^2)
SSR <- sum((fitted(full_model) - mean(mtcars$mpg))^2)
```

```r
SSTotal <- sum((mtcars$mpg - mean(mtcars$mpg))^2)

cat("Predictive model:\n")
```

```
## Predictive model:
```

```r
print(full_model)
```

```
##
## Call:
## lm(formula = mpg ~ cyl + hp + drat + wt, data = mtcars)
##
## Coefficients:
## (Intercept)          cyl           hp         drat           wt
##    34.49588     -0.76229     -0.02089      0.81771     -2.97331
```

```r
cat("\nSSE:", SSE, "\n")
```

```
##
## SSE: 174.3752
```

```r
cat("SSR:", SSR, "\n")
```

```
## SSR: 951.672
```

```r
cat("SSTotal:", SSTotal, "\n")
```

```
## SSTotal: 1126.047
```

```r
## Reduced Model Summary
reduced_model <- lm(mpg ~ cyl + hp, data = mtcars)
summary (reduced_model)
```

```
##
## Call:
## lm(formula = mpg ~ cyl + hp, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.4948 -2.4901 -0.1828  1.9777  7.2934
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 36.90833    2.19080  16.847  < 2e-16 ***
## cyl         -2.26469    0.57589  -3.933  0.00048 ***
## hp          -0.01912    0.01500  -1.275  0.21253
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.173 on 29 degrees of freedom
## Multiple R-squared:  0.7407, Adjusted R-squared:  0.7228
## F-statistic: 41.42 on 2 and 29 DF,  p-value: 3.162e-09
```

```r
## SSR(X3,X4|X1,X2)
SSR_X3X4_given_X1X2 <- sum(residuals(reduced_model)^2) - sum(residuals(full_model)^2)

# partial R-square
R2_full <- summary(full_model)$r.squared
R2_reduced <- summary(reduced_model)$r.squared
R2_partial <- (R2_full - R2_reduced) / (1 - R2_reduced)

cat("SSR(X3, X4|X1, X2):", SSR_X3X4_given_X1X2, "\n")
```

```
## SSR(X3, X4|X1, X2): 117.5993
```

```r
cat("Partial R-square (R2_34|12):", R2_partial, "\n")
```

```
## Partial R-square (R2_34|12): 0.4027725
```

```r
# partial F-test
f_test <- anova(reduced_model, full_model)
print(f_test)
```

```
## Analysis of Variance Table
##
## Model 1: mpg ~ cyl + hp
## Model 2: mpg ~ cyl + hp + drat + wt
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1     29 291.98
## 2     27 174.38  2     117.6 9.1045 0.0009504 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# F-statistic and p-value
f_statistic <- f_test$F[2]
p_value <- f_test$`Pr(>F)`[2]

cat("F-statistic:", f_statistic, "\n")
```

```
## F-statistic: 9.104451
```

```r
cat("p-value:", p_value, "\n")
```

```
## p-value: 0.0009503584
```

```r
# Conclusion
alpha <- 0.05  # Assuming a 5% significance level
if (p_value < alpha) {
  cat("Conclusion: Reject H0. There is significant evidence that at least one of β3 or β4 is not zero.\n")
} else {
```

```
  cat("Conclusion: Fail to reject H0. There is not enough evidence to conclud
e that either β3 or β4 is different from zero.\n")
}
```

```
## Conclusion: Reject H0. There is significant evidence that at least one of
β3 or β4 is not zero. Thus β3, β4 (drat, wt) are significant in predicting
mpg of the cars.
```