Transmission and Mile Per Gallon (MPG)

Cynthia Tang

June 9, 2019

## Summary

The report explored the relationship between transmission types and miles per gallon (MPG). Figure 1 in appendix showed that an automatic transmission trend to be better for MPG, considering other factors. Then, we performed t-test to verify the hypothesis. To quantify the MPG difference between automatic and manual transmissins, we fitted the MPG (outcome) and a set of variables (regressors) in to multiple linar models. We selected tansmission (am), number of cylinders (cyl), gross horsepower (hp), and weight (wt) as repressors and fitted the linear model: .

## Results

The mean of automatic transmission: 24.3923077 miles per gallon.  
The mean of manual transmission: 17.1473684 miles per gallon.

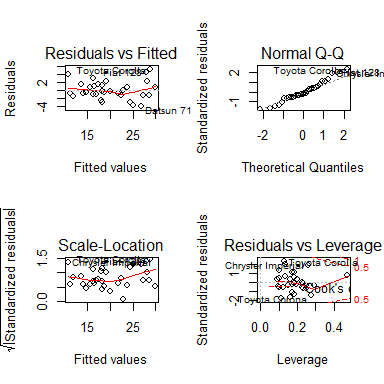
**Model Selection**  
The detailed process of model selection is shown in **Appendix**.  
We set **mpg** as the outcome and **transmission (am)**, **number of cylinders (cyl)**, **gross horsepower (hp)**, and **weight (wt)** as repressors and fitted the linear model: . We assume ~ .  
The coefficients of the linear model are showed below:

## (Intercept) factor(am)manual factor(cyl)6 factor(cyl)8   
## 35.51753528 -1.80921138 -3.03134449 -2.16367532   
## hp wt   
## -0.03210943 -2.49682942

**Interpretion**  
The intercept is interpret as the expected mpg of automatic transmission when there is 4 cylinders, 0 gross horsepower, and none weight.  
The is interpreted as the expected change in mpg comparing those in manual to those in automatic transmission.  
The is interpreted as the expected change in mpg comparing those with 6 cylinders to those with 4 cylinders.  
The is interpreted as the expected change in mpg comparing those with 8 cylinders to those with 4 cylinders.  
The is interpreted as the expected change in mpg for every 1 gross horsepower increase.  
The is interpreted as the expected change in mpg fot every 1000 lbs increase in the weight of cars.

**T-test**  
Assumption: Unequal variance. Detailed test results is shown in **Appendix**. : .  
: .  
The p-value: 6.868191710^{-4}. The type I error = 0.05, then we rejected and concluded that an automatic transmission is better for MPG.

**Quantify the MPG Difference**  
 = 7.2449393.  
Adjustment: = abs() = 1.8092114.

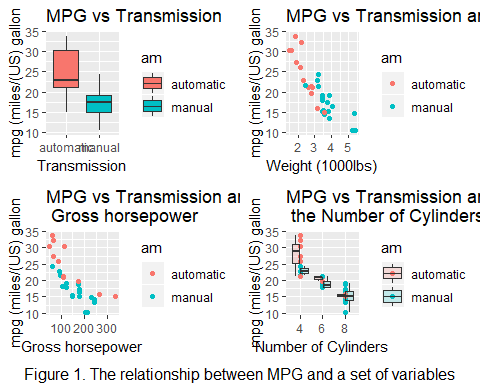
**Diagnostics**  


The mean of residuals is -5.377642810^{-17}, approximating to zero.  
“Residual vs Fitted” is approximately flat, indicating the homoscedasticity of residuals. The Q-Q plot showed that all residuals were approximately stardard normal distributed.

**The uncertainty**  
We assumed other variables not included in the linear model were completely randomized. The error in the linear model followed normal distribution. However, we are uncertain whether there is a better linear model with other combination of diferent regressors.

## Appendix

### Figure



### Model Selection

## Analysis of Variance Table  
##   
## Model 1: mpg ~ factor(am)  
## Model 2: mpg ~ factor(am) + factor(cyl)  
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 30 720.9   
## 2 28 264.5 2 456.4 24.158 8.01e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## Analysis of Variance Table  
##   
## Model 1: mpg ~ factor(am)  
## Model 2: mpg ~ factor(am) + factor(cyl)  
## Model 3: mpg ~ factor(am) + factor(cyl) + disp  
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 30 720.90   
## 2 28 264.50 2 456.40 26.7353 3.956e-07 \*\*\*  
## 3 27 230.46 1 34.04 3.9875 0.05601 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## Analysis of Variance Table  
##   
## Model 1: mpg ~ factor(am)  
## Model 2: mpg ~ factor(am) + factor(cyl)  
## Model 3: mpg ~ factor(am) + factor(cyl) + hp  
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 30 720.9   
## 2 28 264.5 2 456.4 31.2446 9.43e-08 \*\*\*  
## 3 27 197.2 1 67.3 9.2141 0.005266 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## Analysis of Variance Table  
##   
## Model 1: mpg ~ factor(am)  
## Model 2: mpg ~ factor(am) + factor(cyl)  
## Model 3: mpg ~ factor(am) + factor(cyl) + hp  
## Model 4: mpg ~ factor(am) + factor(cyl) + hp + drat  
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 30 720.90   
## 2 28 264.50 2 456.40 30.4123 1.557e-07 \*\*\*  
## 3 27 197.20 1 67.30 8.9686 0.005961 \*\*   
## 4 26 195.09 1 2.11 0.2807 0.600747   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## Analysis of Variance Table  
##   
## Model 1: mpg ~ factor(am)  
## Model 2: mpg ~ factor(am) + factor(cyl)  
## Model 3: mpg ~ factor(am) + factor(cyl) + hp  
## Model 4: mpg ~ factor(am) + factor(cyl) + hp + wt  
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 30 720.90   
## 2 28 264.50 2 456.40 39.286 1.388e-08 \*\*\*  
## 3 27 197.20 1 67.30 11.585 0.002164 \*\*   
## 4 26 151.03 1 46.17 7.949 0.009081 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## Analysis of Variance Table  
##   
## Model 1: mpg ~ factor(am)  
## Model 2: mpg ~ factor(am) + factor(cyl)  
## Model 3: mpg ~ factor(am) + factor(cyl) + hp  
## Model 4: mpg ~ factor(am) + factor(cyl) + hp + wt  
## Model 5: mpg ~ factor(am) + factor(cyl) + hp + wt + qsec  
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 30 720.90   
## 2 28 264.50 2 456.40 39.6232 1.772e-08 \*\*\*  
## 3 27 197.20 1 67.30 11.6849 0.002166 \*\*   
## 4 26 151.03 1 46.17 8.0172 0.009017 \*\*   
## 5 25 143.98 1 7.04 1.2230 0.279293   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## Analysis of Variance Table  
##   
## Model 1: mpg ~ factor(am)  
## Model 2: mpg ~ factor(am) + factor(cyl)  
## Model 3: mpg ~ factor(am) + factor(cyl) + hp  
## Model 4: mpg ~ factor(am) + factor(cyl) + hp + wt  
## Model 5: mpg ~ factor(am) + factor(cyl) + hp + wt + factor(vs)  
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 30 720.90   
## 2 28 264.50 2 456.40 39.7065 1.737e-08 \*\*\*  
## 3 27 197.20 1 67.30 11.7095 0.002146 \*\*   
## 4 26 151.03 1 46.17 8.0341 0.008954 \*\*   
## 5 25 143.68 1 7.35 1.2782 0.268968   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## Analysis of Variance Table  
##   
## Model 1: mpg ~ factor(am)  
## Model 2: mpg ~ factor(am) + factor(cyl)  
## Model 3: mpg ~ factor(am) + factor(cyl) + hp  
## Model 4: mpg ~ factor(am) + factor(cyl) + hp + wt  
## Model 5: mpg ~ factor(am) + factor(cyl) + hp + wt + factor(gear)  
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 30 720.90   
## 2 28 264.50 2 456.40 36.5938 5.143e-08 \*\*\*  
## 3 27 197.20 1 67.30 10.7916 0.003124 \*\*   
## 4 26 151.03 1 46.17 7.4043 0.011916 \*   
## 5 24 149.67 2 1.36 0.1091 0.897096   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## Analysis of Variance Table  
##   
## Model 1: mpg ~ factor(am)  
## Model 2: mpg ~ factor(am) + factor(cyl)  
## Model 3: mpg ~ factor(am) + factor(cyl) + hp  
## Model 4: mpg ~ factor(am) + factor(cyl) + hp + wt  
## Model 5: mpg ~ factor(am) + factor(cyl) + hp + wt + factor(carb)  
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 30 720.90   
## 2 28 264.50 2 456.40 32.9605 3.33e-07 \*\*\*  
## 3 27 197.20 1 67.30 9.7201 0.005206 \*\*   
## 4 26 151.03 1 46.17 6.6691 0.017370 \*   
## 5 21 145.39 5 5.63 0.1627 0.973489   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

### T-test Results

##   
## Welch Two Sample t-test  
##   
## data: mpg by factor(am)  
## t = 3.7671, df = 18.332, p-value = 0.0006868  
## alternative hypothesis: true difference in means is greater than 0  
## 95 percent confidence interval:  
## 3.913256 Inf  
## sample estimates:  
## mean in group automatic mean in group manual   
## 24.39231 17.14737