Automatic Transmission Cars Aren't Worse Than Manual Cars In Terms of MPG

Chi Lam

November 20, 2015

Executive Summary

The Motor Trend Car Road Tests Data was used to determined that automatic transmission cars are not worse than manual transmission cars in terms of fuel economy. It was deteremined that only the *number of cylinders*, and *weight* are significant determinants of a car's *fuel economy*. Forcing *transmission type* into a model doesn't improve the *fuel economy* prediction power of the model. The MPG Difference between an automatic car and a manual car is therefore not quantifiable.

Fitting a Good Model

We fit a linear model that encompasses all the variable, then use an ANOVA table to filter out those variables that are not useful. To contrast the result, we will later on try to force *transmission type* into our model. The following ANOVA table shows that only the *number of cylinders*(cyl), and the *weight*(wt) are significant predictors for the cars' MPG at a two-sided 95% confidence level using the dataset. *Displacement*(disp) fails the two-sided 95% confidence level test by a small margin.

```
## Analysis of Variance Table
##
## Response: mpg
##           Df Sum Sq Mean Sq  F value    Pr(>F)
## cyl        1 817.71  817.71 116.4245 5.034e-10 ***
## disp       1  37.59   37.59   5.3526  0.030911 *
## hp         1   9.37    9.37   1.3342  0.261031
## drat       1  16.47   16.47   2.3446  0.140644
## wt         1  77.48   77.48  11.0309  0.003244 **
## qsec       1   3.95    3.95   0.5623  0.461656
## vs         1   0.13    0.13   0.0185  0.893173
## am         1  14.47   14.47   2.0608  0.165858
## gear       1   0.97    0.97   0.1384  0.713653
## carb       1   0.41    0.41   0.0579  0.812179
## Residuals 21 147.49    7.02
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Therefore, the best model balancing between bias and over-modeling is:

`mpg = B0 + B1*cyl + B2*wt`

Note that the *transmission type*(am) is not a factor of the model.

Best Fit Model Coefficients

```
##        B0        B1        B2
## 39.686261 -1.507795 -3.190972
```

The above are the estimated coefficients. Hence, the average *fuel economy*(mpg) of the cars is 39.6862615 mpg; and for each unit increase in *number of cylinder*(cyl), a change of -1.507795 mpg of *fuel economy*(mpg) is expected; and for each 1000 lbs increase in *weight*(wt), a change of -3.1909721 mpg of *fuel economy*(mpg) is expected.

Uncertainties in the Coefficients

The 95% confidence intervals for the coefficients are calculated as the following:

```
##    lowerBound upperBound
## B0  36.178725 43.1937976
## B1  -2.355928 -0.6596622
## B2  -4.739020 -1.6429245
```

Note that none of intervals contain 0; therefore, this is an acceptance model. Residual plots are also available in the appendix. The plots show that the residuals are normal and random. These two facts give us confidnece to this best fit model.

Forcing Transmission Type As a Predictor

The *Transmission type*(am) is added to the best fit model to investigate further if *transmission type* can affect *fuel economy.* The following ANOVA table compares the two models.

```
## Analysis of Variance Table
##
## Model 1: mpg ~ cyl + wt
## Model 2: mpg ~ cyl + wt + am
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     29 191.17
## 2     28 191.05  1   0.12491 0.0183 0.8933
```

Note that the p value for the new model is huge(0.89) compared to our threshold of 0.025 for a 95% confidence level. This highly suggests that the additional variable is not necessary. The 95% confidence interval for the slope coefficient of *transmission type* follows:

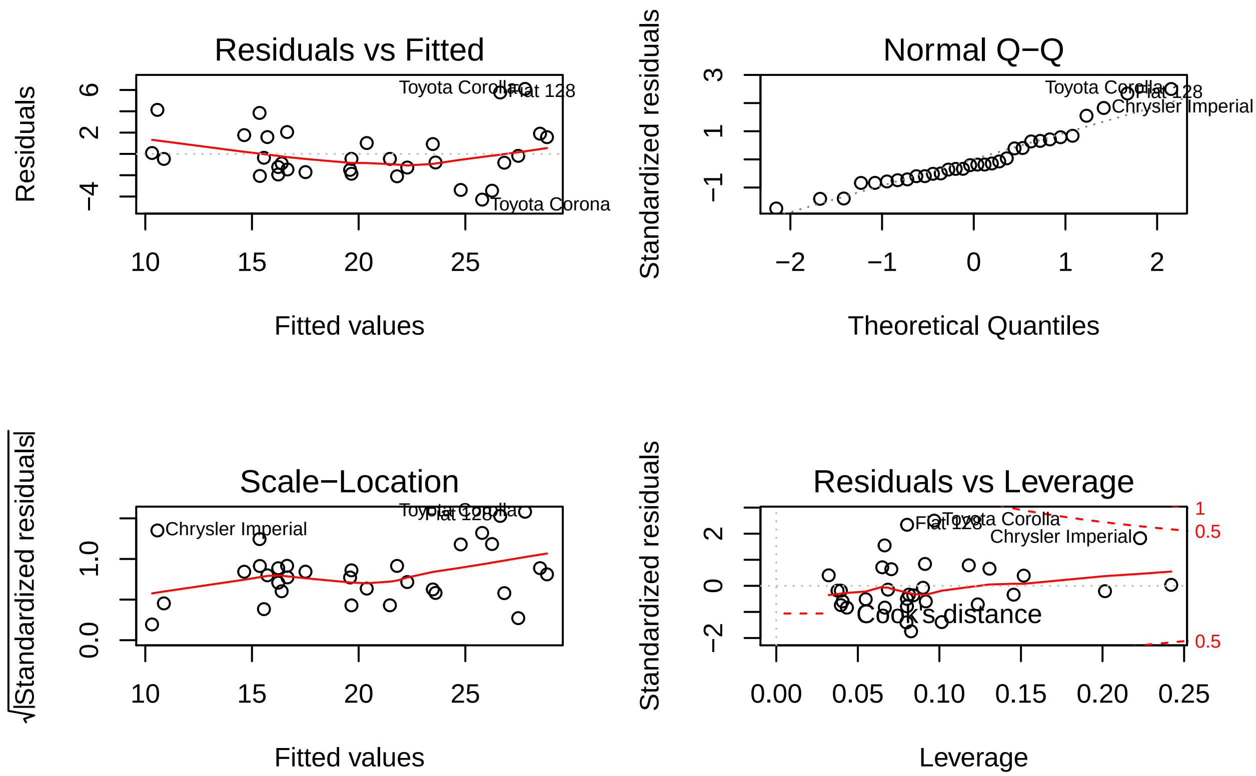
```
## [1] -2.495555  2.848541
```

Note that the interval contains 0, therefore further disproving the possibility that *transmission type* being a good predictor for a car's *fuel economy*(mpg).

Conclusion

Using an ANOVA table, it was deteremined that only the *number of cylinders*, and *weight* are significant determinants of a car's *fuel economy*, not *transmission type*. Forcing *transmission type* into a model doesn't improve the *fuel economy* prediction power of the model. Therefore, we can conclude that automatic transmission cars are not statistically different from manual transmission cars in terms of *fuel economy.* The MPG Difference between an automatic car and a manual car is therefore not quantifiable.

Appendix: Residual Plots of the Best Fit Model



The residual plots look normal, and there doesn't seem to be a pattern in the residual; therefore the best fit model is an acceptable model.

Appendix: Code used

Fitting a Good Model

```
cars <- mtcars
cars$am <- factor(cars$am)
cars$vs <- factor(cars$vs)
carAllFit <- lm(mpg ~ ., data=cars)
anova(carAllFit)
```

Best Fit Model Coefficients

```
carBestFit <- lm(mpg ~ cyl + wt, data=cars)
bestFitCoeff <- summary(carBestFit)$coefficients
row.names( bestFitCoeff ) <- c("B0", "B1", "B2")
bestFitCoeffEst <- bestFitCoeff[,1]
bestFitCoeffEst
```

Uncertainties in the Coefficients

```
bestFitCoeffConInt <- qt(.975, df = carBestFit$df)* bestFitCoeff[,2]
bestFitCoeffLB <- bestFitCoeffEst - bestFitCoeffConInt
bestFitCoeffUB <- bestFitCoeffEst + bestFitCoeffConInt
bestFitCoeffBounds <- data.frame(lowerBound=bestFitCoeffLB, upperBound=bestFitCoeffUB)
bestFitCoeffBounds
```

Forcing Transmission Type As a Predictor

```
cylWtAmFit <- lm(mpg ~ cyl + wt + am, data=cars)
anova(carBestFit, cylWtAmFit)
```

```
cylWtAmFitCoeff <- summary(cylWtAmFit)$coefficients
amCoeffBounds <- cylWtAmFitCoeff[4,1]+ c(-1, 1) * qt(.975, df = cylWtAmFit$df)*
    cylWtAmFitCoeff[4,2]
amCoeffBounds
```

Residual Plots of the Best Fit Model

```
par(mfrow = c(2, 2))
plot(carBestFit)
```