Talking Transmissions-Impacts on Fuel Economy

Larkin O'Hern

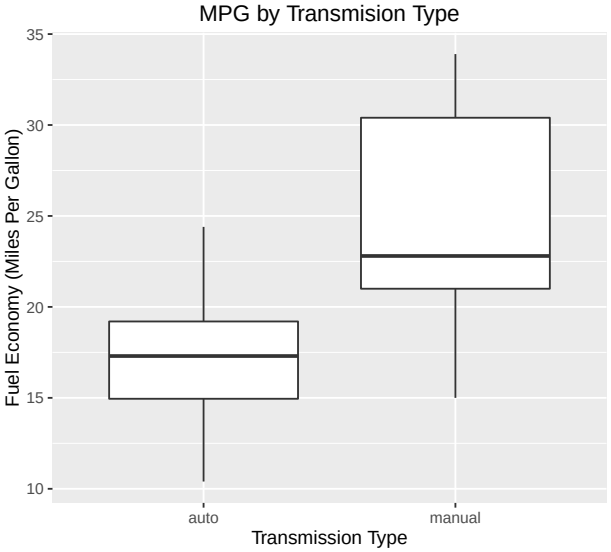
July 13, 2016

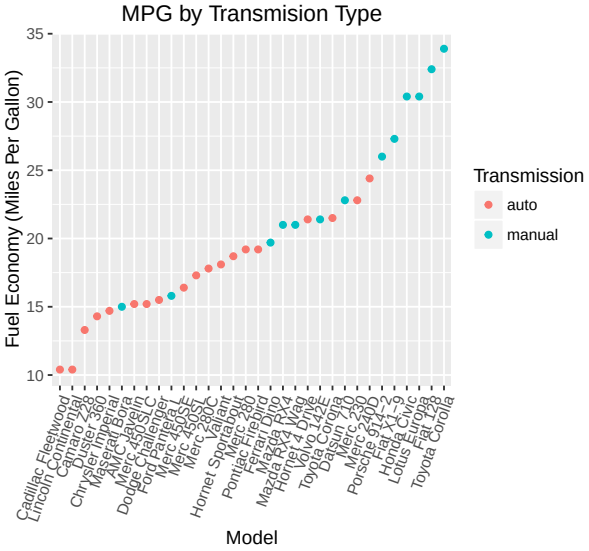
Executive Summary

This analysis will explore the relationship between fuel economy-measured as miles per gallon-and vehicle transmission type-either automatic or manual to determine:
1. Is an automatic or manual transmission better for MPG?
2. The quantitative MPG difference between automatic and manual transmissions.

Exploratory Data Analysis

The mtcars data contains observations on 11 design and performance characteristics of 32 cars drawn from the 1974 Motor Trend US magazine. I will focus primarily on transmission type (manual vs automatic) and fuel economy (in miles per gallon). The cars in the data set are almost evenly split between manual (13) and automatic (19) giving plenty of observations for each. Fuel economy ranges from 10.4-33.9 MPG with a mean 20.09 and median 19.02. If we condition on transmission type stark difference emerge.





From the histogram we see that cars with manual transmissions have higher average fuel economy and great variability. From the scatter plot we see that almost all f the low fuel economy cars are automatics while most of the high fuel economy cars are manual.

We can use a quick t test to determine whether the difference in means is statistically significant.

Zero is not included in our confidence interval (-11.2801944, -3.2096842) and our p-value (0.0013736) is below .05 so transmission type appears to make a statistically significant difference in fuel economy.

Model Selection

To quantify the effect of transmission type on fuel economy we will fit a series of linear models of increasing complexity.

The simplest model examines transmission type as the only predictor of mpg.

```
##
## Call:
## lm(formula = mpg ~ am, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.3923 -3.0923 -0.2974  3.2439  9.5077
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   17.147      1.125  15.247 1.13e-15 ***
## ammanual       7.245      1.764   4.106 0.000285 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.902 on 30 degrees of freedom
## Multiple R-squared:  0.3598, Adjusted R-squared:  0.3385
## F-statistic: 16.86 on 1 and 30 DF,  p-value: 0.000285
```

```
##                2.5 %   97.5 %
## (Intercept) 14.85062 19.44411
## ammanual     3.64151 10.84837
```

This basic model confirms that transmission type is a significant predictor and the coefficient for transmission type indicates an increase of 7.25 MPG for vehicles with manual transmissions versus those with automatic transmission-all else being equal. Exanding our point prediction to encompasses a 95% confidence interval gives us an expected increase of 3.4 to 10.84 mpg in cars with manual transmissions.

However, transmission type explain only about 35% of thevariance in MPG-so there are likly other signficant variables that could add explanatory power to our model.

Here I fit five additional models each with one additional predictor or interaction between transmissin type and predictor (nested models) and compare them using ANOVA.

```
## Analysis of Variance Table
##
## Model 1: mpg ~ am
## Model 2: mpg ~ am + wt
## Model 3: mpg ~ am + wt + am:wt
## Model 4: mpg ~ am + wt + hp + am:wt
## Model 5: mpg ~ am + wt + disp + hp + am:wt
## Model 6: mpg ~ am + wt + disp + hp + cyl + am:wt
## Model 7: mpg ~ am + wt + disp + hp + cyl + gear + am:wt
##   Res.Df    RSS Df Sum of Sq       F    Pr(>F)
## 1     30 720.90
## 2     29 278.32  1    442.58 79.8463  4.22e-09 ***
## 3     28 188.01  1     90.31 16.2934 0.0004805 ***
## 4     27 146.85  1     41.16  7.4262 0.0118058 *
## 5     26 144.22  1      2.63  0.4736 0.4979279
## 6     25 133.83  1     10.39  1.8738 0.1837195
## 7     24 133.03  1      0.81  0.1452 0.7064766
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Vehicle weight and the interaction of weight and transmission type add predictive power-the othervariables do not.

I've re-fit the model using the signficant predictors togenerate a new estimate of the impact oftransmission type after accounting for the other key variables.

```
##
## Call:
## lm(formula = mpg ~ am + wt + am:wt, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.6004 -1.5446 -0.5325  0.9012  6.0909
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  31.4161     3.0201  10.402 4.00e-11 ***
## ammanual     14.8784     4.2640   3.489  0.00162 **
## wt           -3.7859     0.7856  -4.819 4.55e-05 ***
## ammanual:wt  -5.2984     1.4447  -3.667  0.00102 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.591 on 28 degrees of freedom
## Multiple R-squared:  0.833,  Adjusted R-squared:  0.8151
## F-statistic: 46.57 on 3 and 28 DF,  p-value: 5.209e-11
```

```
##                 2.5 %    97.5 %
## (Intercept) 25.229642 37.602469
## ammanual     6.143928 23.612917
## wt          -5.395234 -2.176581
## ammanual:wt -8.257693 -2.339028
```