MPG Comparison of Automatic and Manual Transmissions

D.Barger, Coursera REMODS-008 Course Project

Friday, November 23, 2014

Executive Summary

Generally speaking, most drivers believe automobiles with manual transmission have better miles per gallon than automatic transmission. Based on the following analysis using data collected by "Motor Trend" magazine in 1974, manual transmission automobiles get 2.084 MPG more thant automatic transmission for every mile driven.

Exploratory Analysis

Based on the graphs in the Appendix ("Exploratory Analysis"), the data follows a normal distribution. On average, automobiles with manual transmission get an additional 7.24 miles per gallon with an interquartile range of 9.4 miles per gallon.

```
##            am    mpg
## 1 Automatic 17.15
## 2    Manual 24.39
```

Hypothesis Testing: Is there a Significant Difference?

The initial results indicate automobiles with manual transmission provide better MPG, but is the difference significant? Hypothesis testing will determine significance and whether additional analysis is beneficial.

```
a <- mtcars[mtcars$am=="Automatic",]
m <- mtcars[mtcars$am=="Manual",]
t.test(a$mpg,m$mpg)

##
##  Welch Two Sample t-test
##
## data:  a$mpg and m$mpg
## t = -3.767, df = 18.33, p-value = 0.001374
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.28  -3.21
## sample estimates:
## mean of x mean of y
##     17.15     24.39
```

The P-Value is 0.001374 and the difference in means is not zero. There is a significant difference between manual and automatic transmissions as it relates to miles per gallon (MPG).

Model Selection

There are numerous categorical variables in the mtcars dataset. To ensure model results are not skewed, selection of appropriate predicators is important. It seems horsepower (hp) and automobile weight (wt) are the best predicators for our model.

```
data(mtcars)
cor(mtcars)[1,]

##      mpg     cyl    disp      hp    drat      wt    qsec      vs      am
##   1.0000 -0.8522 -0.8476 -0.7762  0.6812 -0.8677  0.4187  0.6640  0.5998
##     gear    carb
##   0.4803 -0.5509
```

Regression Modeling

Before conducting any multivariate regression analysis, let's fit a simple regression between mpg and transmission types (am)

```
mtcars$am <- factor(mtcars$am, labels=c("Automatic","Manual"))
fit <- lm(mpg ~ am, data=mtcars)
summary(fit)$coef

##               Estimate Std. Error t value  Pr(>|t|)
## (Intercept)    17.147       1.125  15.247 1.134e-15
## amManual        7.245       1.764   4.106 2.850e-04
```

THe relationship between variables am and mpg only explain around 36% of variance (Multiple R-squared). Additional variables are needed int the model to gain more accuracy.

```
fit <- lm(mpg ~ am + wt + hp, data=mtcars)
summary(fit)$coef

##               Estimate Std. Error t value  Pr(>|t|)
## (Intercept) 34.00288    2.642659  12.867 2.824e-13
## amManual     2.08371    1.376420   1.514 1.413e-01
## wt          -2.87858    0.904971  -3.181 3.574e-03
## hp          -0.03748    0.009605  -3.902 5.464e-04
```
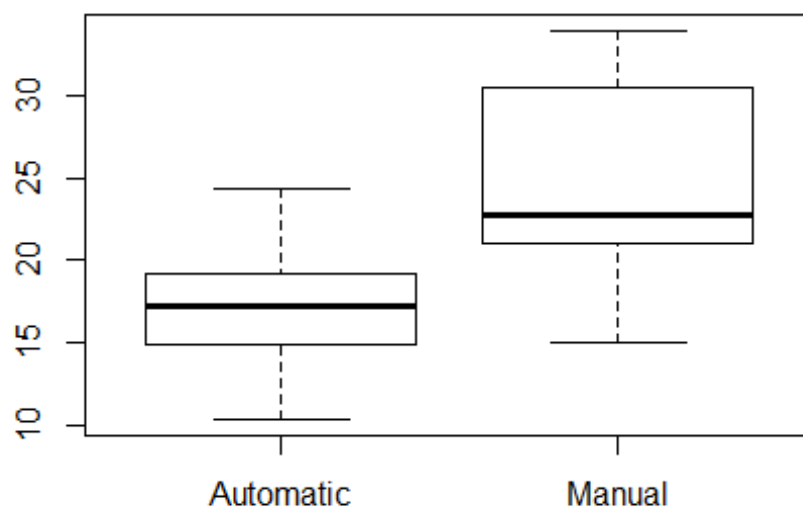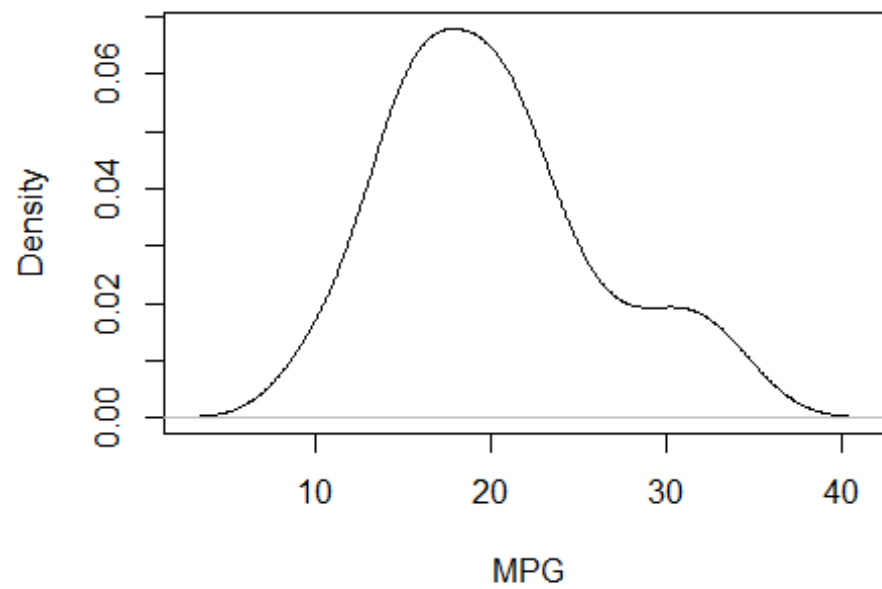
Conclusion

Adding hp and wt to model as predicators explained 84% of the model variance. As a result, the model proved manual transmission get 2.085 miles per gallon than automobiles with automatic transmission.

Appendix

Exploratory Analysis Graphs

MPG Density Plot

Regression Modeling Graphs

Residuals vs Fitted

Residuals vs Fitted plot. Points labeled: Toyota Corolla, Fiat 128, Chrysler Imperial. X-axis: Fitted values, 10 to 30. Y-axis: Residuals, -4 to 6.

lm(mpg ~ am + wt + hp)

Normal Q-Q

Normal Q-Q plot. Points labeled: Toyota Corolla, Fiat 128, Chrysler Imperial. X-axis: Theoretical Quantiles, -2 to 2. Y-axis: Standardized residuals, -1 to 2.

lm(mpg ~ am + wt + hp)

Scale-Location

√|Standardized residuals|

Chrysler Imperial

Toyota Corolla

Fiat 128

Fitted values
lm(mpg ~ am + wt + hp)

Residuals vs Leverage

Standardized residuals

Toyota Corolla

Fiat 128
Chrysler Imperial

Cook's distance

1

0.5

0.5

Leverage
lm(mpg ~ am + wt + hp)