

Cars MPG Analysis

Amade A.

December 25, 2015

Overview

Research objective

The objective of this paper is answering the question whether automatic transmission is better for MPG compared to manual transmission. In addition, the researcher also aims to quantify the difference in MPG between cars with different types of transmission.

Executive summary

This paper analyzes the influence of various car characteristics on fuel economy (measured as miles per gallon of fuel).

Exploratory Data Analysis

Dataset

Dataset used for the analysis is **mtcars** dataset, preloaded in R in the *datasets* package. Data originates from the 1974 Motor Trend US magazine, measuring fuel consumption across 10 automobile design aspects for 32 automobiles.

Among the 10 variables, 5 (cyl, vs, am, gear, carb) can be characterized as discrete (factor) variables and remaining 5 (disp, hp, drat, wt, qsec), continuous (numeric) variables. Fuel consumption is represented by the continuous *mpg* variable.

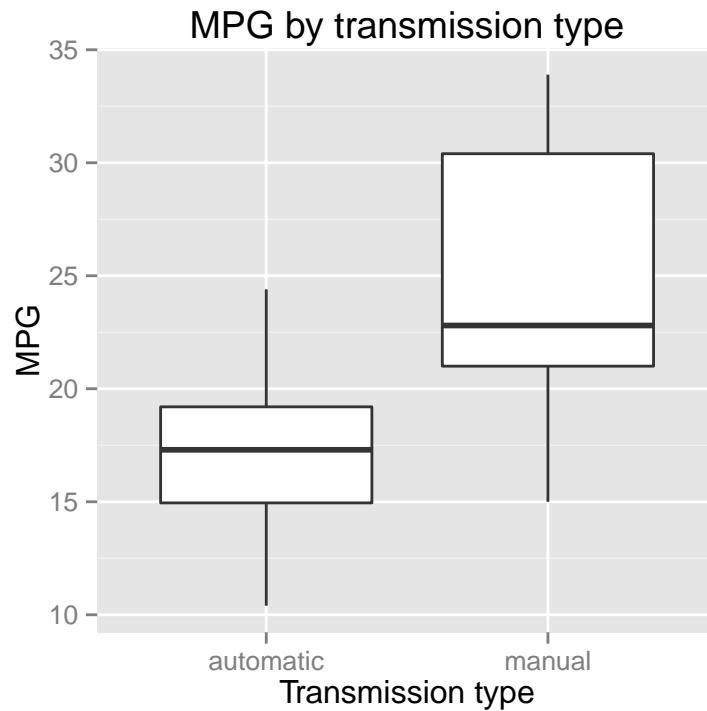
For easier interpretation, variable specifying the transmission type - *am* was recoded, so that the factor levels have straightforward interpretation.

```
library(dplyr, warn.conflicts = FALSE)
mtcars <- mtcars %>%
mutate(am = factor(am, levels = c(0,1), labels = c("automatic", "manual")))
```

MPG by transmission type

In total, there were 19 cars with automatic transmission and 13 with manual. The following graph shows the differences in measured fuel efficiency among cars with different transmission types.

```
library(ggplot2)
p <- ggplot(mtcars, aes(factor(am), mpg))
p <- p + geom_boxplot()
p <- p + ggtitle("MPG by transmission type") + xlab("Transmission type") + ylab("MPG")
p
```



It can be easily seen, that the fuel efficiency is not the same among the cars with different transmission types.
Formal modelling

Simple model (automatic/manual only)

Inclusion of possible confounding variables

Best model selection

Here perform a likelihood ratio test

Uncertainty measures

Analysis of residuals

Here plot the residuals

Confidence intervals for predictor variables

Specify here the confidence intervals for those variables

Conclusions and interpretation

Interpret the outcome of the model