Motor Trends Analysis

wdewit

October 8, 2016

Executive Summary

In this analysis we investigated the relationship between fuel consumption and the transmission type. We had information on both variables as well as a set of 9 other car characteristics at our disposal for 32 car brands.

Exploratory Data Analysis

Since fuel consumption (mpg) is a continous variable we will estimate a linear regression model. One of the assumptions of a linear model is that the dependent variable is normally distributed.

Comparing the density of mpg with the normal density we see mpg is somewhat skewed to the right (fig 1 - Appendix). However the deviation from normality is not too strong and probably caused by our rather small sample size. Given that least squares estimation is quite robust against violation of the normality assumption, we continue fitting a linear regression model by means of least squares.

Next we'll have a first glance whether there does seem to be a relationship between transmission type and fuel consumption.

```
tapply(mtcars$mpg,mtcars$am,mean)
```

```
##        0        1
## 17.14737 24.39231
```

We observe that average fuel consumption with an automatic transmission is 17.14 mpg, which is nearly one third less than the average fuel consumption with a manual transmission. From a boxplot (fig 2 - Appendix) we also notice that the fuel consumption with a manual transmission is far more variable than with an automatic transmission, especially skewed towards higher values of mpg.

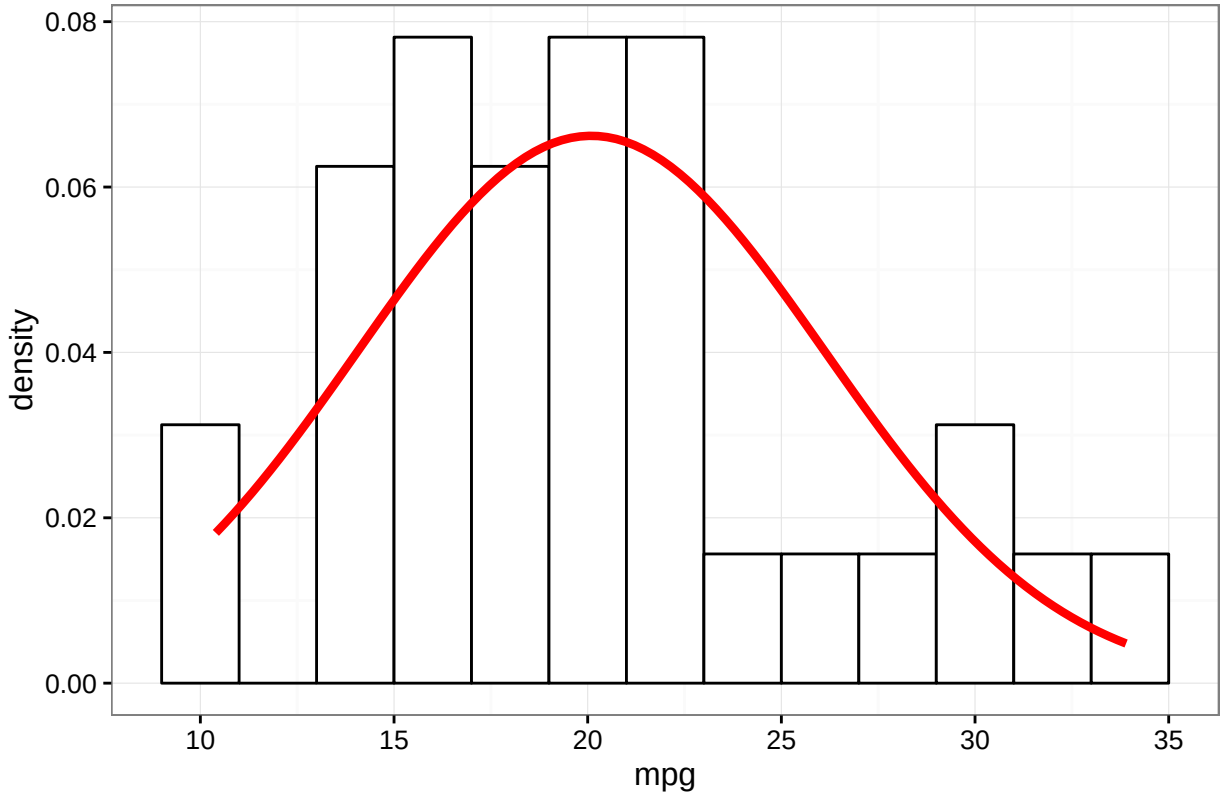
Formal modelling

We now want to confirm whether there is a statistically significant difference between fuel consumption of automatic versus manual transmission types.

Appendix:

```
library(ggplot2)
g <- ggplot(data=mtcars,aes(x = mpg))+ labs(title="Fig 1: Checking normality of our dependent variable")
g <- g + geom_histogram(aes(y = ..density..), fill = "white", binwidth=2, colour = "black")
g + stat_function(fun=dnorm,colour="red",
                  size=1.5, args=list(mean=mean(mtcars$mpg), sd=sd(mtcars$mpg)))+theme_bw()
```

Fig 1: Checking normality of our dependent variable




```
boxplot(mtcars$mpg~mtcars$am,xlab="transmission type (0=automatic 1=manual)", ylab="fuel consumption (mp
```

Fig 2: Comparing distribution of mpg by transmission type