

Regmod

Steven Zwart

Saturday, November 22, 2014

This document describes an analysis of the mtcars table using R.

The following 2 questions will be answered:

“Is an automatic or manual transmission better for MPG” “Quantify the MPG difference between automatic and manual transmissions”

```
summary(mtcars)
```

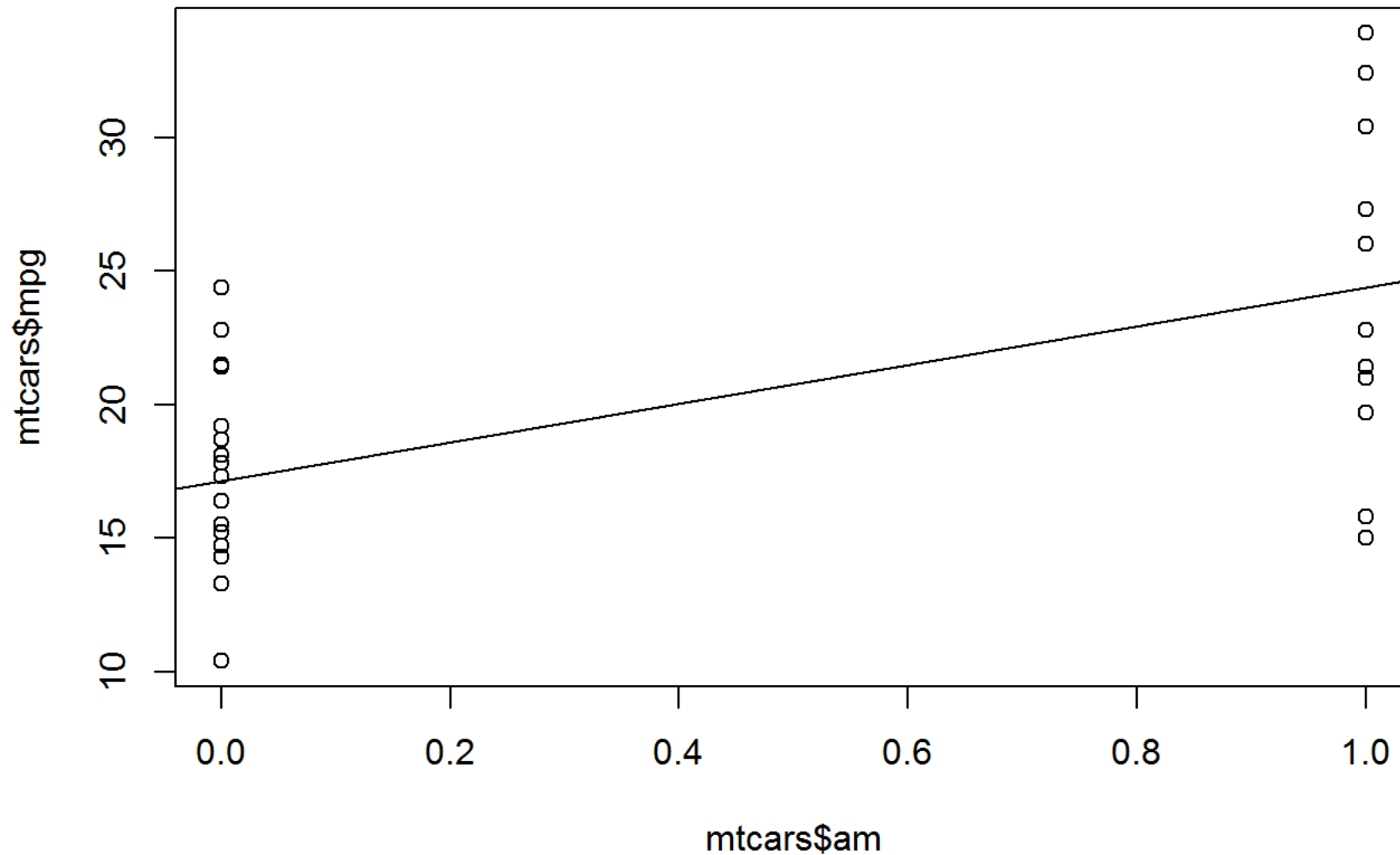
```
##           mpg           cyl           disp           hp
##  Min.      :10.4   Min.      :4.00   Min.      : 71.1   Min.      : 52.0
##  1st Qu.:15.4   1st Qu.:4.00   1st Qu.:120.8   1st Qu.: 96.5
##  Median :19.2   Median :6.00   Median :196.3   Median :123.0
##  Mean     :20.1   Mean     :6.19   Mean     :230.7   Mean     :146.7
##  3rd Qu.:22.8   3rd Qu.:8.00   3rd Qu.:326.0   3rd Qu.:180.0
##  Max.     :33.9   Max.     :8.00   Max.     :472.0   Max.     :335.0
##           drat           wt           qsec           vs
##  Min.      :2.76   Min.      :1.51   Min.      :14.5   Min.      :0.000
##  1st Qu.:3.08   1st Qu.:2.58   1st Qu.:16.9   1st Qu.:0.000
##  Median :3.69   Median :3.33   Median :17.7   Median :0.000
##  Mean     :3.60   Mean     :3.22   Mean     :17.8   Mean     :0.438
##  3rd Qu.:3.92   3rd Qu.:3.61   3rd Qu.:18.9   3rd Qu.:1.000
##  Max.     :4.93   Max.     :5.42   Max.     :22.9   Max.     :1.000
```

##	am	gear	carb
##	Min. :0.000	Min. :3.00	Min. :1.00
##	1st Qu.:0.000	1st Qu.:3.00	1st Qu.:2.00
##	Median :0.000	Median :4.00	Median :2.00
##	Mean :0.406	Mean :3.69	Mean :2.81
##	3rd Qu.:1.000	3rd Qu.:4.00	3rd Qu.:4.00
##	Max. :1.000	Max. :5.00	Max. :8.00

```
corvec <- c()
for (i in 1:11) {
  if (i!=9) {
    corvec <- c(corvec,cor(mtcars[,i], mtcars$am))
  }
}
```

The correlation between transmission and mpg is 0.600 meaning that manually transmitted cars appear to have a higher mpg. However, there is a relatively high correlation between am and cyl, am and disp, am and drat, am and wt, am and gear, so these may be confounding variables.

```
linmod <- lm(mtcars$mpg~mtcars$am)
plot(mtcars$mpg ~mtcars$am)
abline(linmod)
```



Now let's take possible confounding factors into consideration:

```
linmod <- lm(mtcars$mpg~mtcars$am+mtcars$cyl+mtcars$disp+mtcars$drat+mtcars$wt+mtcars$gear)
summary(linmod)
```

```
##
## Call:
## lm(formula = mtcars$mpg ~ mtcars$am + mtcars$cyl + mtcars$disp +
##      mtcars$drat + mtcars$wt + mtcars$gear)
##
## Residuals:
##      Min        1Q    Median        3Q        Max
## -4.636 -1.446 -0.466  1.674  5.604
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  42.84931    7.74036   5.54  9.4e-06 ***
## mtcars$am      1.18158    1.87872   0.63  0.5351
## mtcars$cyl    -1.74082    0.65474  -2.66  0.0135 *
## mtcars$disp    0.00571    0.01248   0.46  0.6515
## mtcars$drat    0.33110    1.61860   0.20  0.8396
## mtcars$wt     -3.42536    1.22605  -2.79  0.0099 **
## mtcars$gear   -1.07237    1.15316  -0.93  0.3613
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.7 on 25 degrees of freedom
## Multiple R-squared:  0.838, Adjusted R-squared:  0.799
## F-statistic: 21.6 on 6 and 25 DF, p-value: 9.1e-09
```

In this case we get a t-value of 0.629, and a p_value of 0.54, so by no means significant. Using only cylinders and weight as confounding factors gives the following:

```
linmod <- lm(mtcars$mpg~mtcars$am+mtcars$cyl+mtcars$wt)
summary(linmod)
```

```
##
## Call:
## lm(formula = mtcars$mpg ~ mtcars$am + mtcars$cyl + mtcars$wt)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4.173 -1.534 -0.539  1.586  6.081
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   39.418     2.641   14.92  7.4e-15 ***
## mtcars$am      0.176     1.304    0.14  0.8933
## mtcars$cyl    -1.510     0.422   -3.58  0.0013 **
## mtcars$wt     -3.125     0.911   -3.43  0.0019 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.61 on 28 degrees of freedom
## Multiple R-squared:  0.83,    Adjusted R-squared:  0.812
## F-statistic: 45.7 on 3 and 28 DF,  p-value: 6.51e-11
```

So concluding we can give the following answers 1. We have no evidence that automatic or manual transmission have influence on the mpg 2. The slope estimate is that for changing from automatic to manual (while adjusting for cylinder and weight) the mpg increases by 0.1765. This result however is not significant (p value of 0.89334)

