Do Manual Transmissions Improve Gas Mileage?

Peter Putz

November 21, 2014

In this paper[1] we investigate whether an automatic or manual transmission is better for gas mileage and if so, how big the difference is. Unfortunately, the findings indicate that the available data is not sufficient for a conclusive answer.

[1] This paper is the response to an assignment for the Coursera course "Regression Models" by Brian Caffo, Jeff Leek, and Roger Peng.

THE DATASET for this analysis is "mtcars" from the "R datasets package". It was extracted from the 1974 *Motor Trend* US magazine, and comprises fuel consumption and 10 aspects of automobile design and performance for 32 automobiles (1973–74 models).

An *exploratory analysis* suggests that the average mileage of cars with manual transmissions is 24.4 mpg compared to 17.1 mpg for automatic transmissions. A two sample t-test shows that this difference is statistically significant (p = 0.001).

In the following sections we will investigate whether this finding is confounded by other factors and if so, what impact, if any, the transmission type actually has onto a car's fuel efficiency.

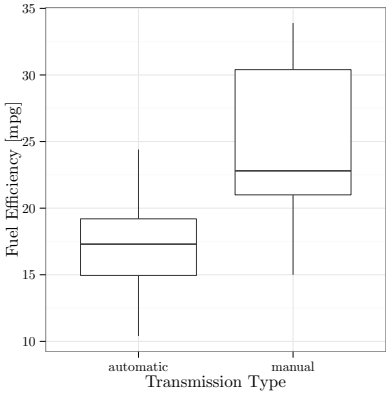


Figure 1: Mileage for automatic vs. manual transmissions

OUR INITIAL MODEL is guided by physics: "To the first order, fuel economy is essentially inversely proportional to auto mass (...)"[2].

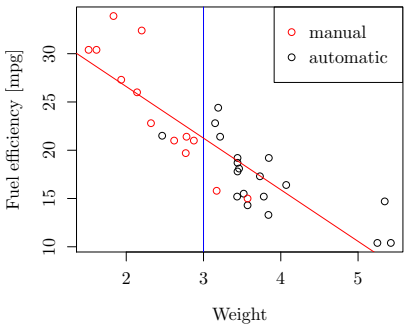
[2] Hafemeister, David. 2013. Physics of Societal Issues. 2nd Edition. Springer, New York, p. 513.

```
model1 <- lm(mpg ~ wt, mtcars)
```

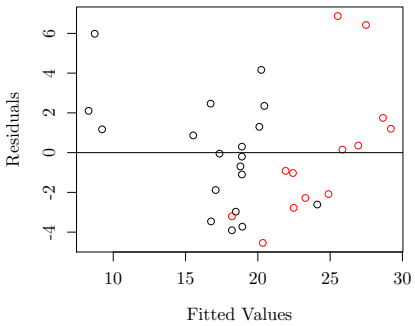
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	37.2851	1.8776	19.86	0.0000
wt	-5.3445	0.5591	-9.56	0.0000

$R^2_{adj} = 0.7446$

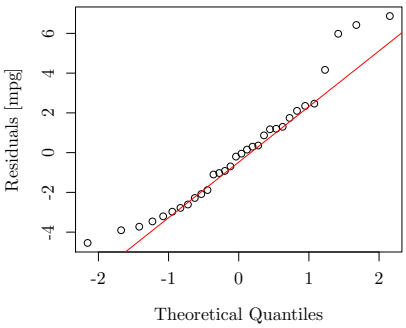
(2a) mpg = 37.3 - 5.35wt



(2b) Residual Plot



(2c) Q-Q Plot



Model 1 explains 74.5% of the total variance of mpg and it suggests that 1,000 lbs. increase in weight reduces the fuel efficiency by 5.34

mpg (±1.14 mpg). The formal conditions for this interference are met: Linearity (figure 2a), constant variability of residuals (2b), and nearly normal residuals (2c). In model 2 we add the transmission type "am".

```
model2 <- lm(mpg ~ wt + factor(am), mtcars)
```

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	37.3216	3.0546	12.22	0.0000
wt	-5.3528	0.7882	-6.79	0.0000
factor(am)1	-0.0236	1.5456	-0.02	0.9879

$R^2_{adj} = 0.7358$

Clearly, model 2 does not offer a better fit than model 1. In fact, R^2_{adj} declines from 74.5% to 73.6%. In addition, the p-value for the dummy variable *am* is extremely high ($p = 0.988$). Therefore, we cannot reject the null hypothesis that the coefficient for *am* is zero. This suggests, that the transmission type has no influence on the fuel economy.

However, there is another problem which becomes apparent in figure 2a. In our data set, the transmission type is highly dependent on the weight of a car. The blue line in figure 2a separates cars with less than 3,000 lbs. from cars with more than 3,000 lbs. It highlights, that with a few exceptions lightweight cars in our dataset happen to have a manual transmission while heavier vehicles have automatic transmissions. This suggests that *there is not sufficient data to inform the relationship between transmission type and fuel economy.*

This result does not change if we try to improve the overall fit of the model. For model 3 we add the *quarter mile time* (qsec) for theoretical reasons (qsec is a proxy for acceleration) and because it is minimally correlated with weight ($cor = -0.175$).

```
model3 <- lm(mpg ~ wt + qsec, mtcars)
```

Compared to the initial model 1, R^2_{adj} increases from 74.5% to 81.4% indicating that it is a better fit. In model 4 we add the transmission type back in:

```
model4 <- lm(mpg ~ wt + qsec + factor(am), mtcars)
```

While R^2_{adj} further increases to 83.4% the p-values for the transmission coefficients (intercept and factor(am)1) are high. This indicates that adding the transmission type does not improve model 3.

In summary,[3] based on the dataset "mtcars", we could not show whether or not the transmission type has any impact onto the fuel economy of a car. To answer this question, further data is needed that includes both transmission types in lightweight and heavy cars.

	Estimate	Pr(>|t|)
(Intercept)	19.75	0.00
wt	-5.05	0.00
qsec	0.93	0.00

Table 1: Model 3

	Estimate	Pr(>|t|)
(Intercept)	9.62	0.18
wt	-3.92	0.00
qsec	1.23	0.00
factor(am)1	2.94	0.05

Table 2: Model 4

[3] This report was created with knitr. The original .Rmd file is attached to this pdf file.