MPG Comparisons between Manual and Automatic Transmissions

J.E. Turcotte

May 17, 2016

Executive Summary

In this brief study, we must explore the effect that automatic versus manual transmissions has over miles-per-gallon fuel efficiency among a sample group of 1974 model automobiles. First we looked at the data itself, in brief. This we followed up with a simple examination of miles per gallon as predicted by transmission type, but included indicators that suggested that the real factors might be elsewhere. Following this, we fit a simple model and compare it to a multivariate model that included measured for transmission type, holding cylinders, horsepower and weight converted constant. For further illustration, we view each regressor separately.

Assumptions

- The field 'qsec' (or, from a stop, how many seconds it took for the vehicle to reach a quarter mile's distance) is likely a function of the car's power; i.e., a result, like miles per gallon, of the vehicle's configuration rather than a potential confounding cause OF miles per gallon. For that reason, I'm excluding it in any modelling of miles per gallon per vehicle.
- Given the yes-or-no nature of whether or not a car is manufactured with or without automatic transmission, this data has been categorized, so named within the data, and is likely to be modelled using binomial logistic regression.
- This study assumes that the engine power of each given vehicle proportional to the vehicle's weight. Rather, that the mass of the vehicle requires a matching size engine, and that a bigger engine will also contribute to the weight of the vehicle. Thus, a disproportionately designed vehicle might show up as an outlier.

Examining the Data

Looking at *Figure 1*, we get a very quick sense of what the data that we ahve to work with looks like. Given that not all these columns (already exlcuding 'vs' and 'qsec') are easily read or in quite the right format, the first step I've taken here is to rename the columns and change the transmission data to a named factor.

Following this, *Figure 2* gives us a brief summarization of the better-named data that remains, excluding a few extra columns for brevity.

A very simple *Figure 3* binomial logistic regression line drawn purely comparing miles per gallon (MPG) to whether (1) or not (0) a car had a manual transition, there does APPEAR to be a strong differentiation, BUT… we've yet to examine any other details of these cars. Further, giving the data points characteristics (size in tonnage and color, re: number of cylinders), the subpatterning suggests that either or both are significant factors… rather, that they are consistent with the same curve that we might have assigned to choice of transmission alone.

Comparing coefficients between a simple and strict linear regression, in *Figure 4* between Miles per Gallon (MPG) and whether or not any given car is a manual transmission versus a wider net cast, in *Figure 5* with several other regressors (being manual transmission, cylinders, horsepower, and tonnage (converted from the original measure of half-tons)), it becomes rather apparent that the use of transmission is not the prevailing factor.

Hypthotheses (H_0 **v.** H_a)

Given the linear regression:

$$Y = 1 + \beta_i^{tr} X_i + \beta_i^{cyl} X_i + \beta_i^{wt} X_i + \epsilon_i$$

…This paper shall **assume** that the choice of transmission is not a significantly deterministic factor in a vehicle's miles-per-gallon. Rather, in H_0, that $\beta^{tr} < \beta^{wt} + \beta^{cyl}$ rather than in H_a where $\beta^{tr} >= \beta^{wt} + \beta^{cyl}$.

Conclusion

Notice that when considering only the type of transmission, we appear to gain about 7.25 miles per gallon by going with a manual, but when regressing for number of cylinders, 0.75mpg is removed per cylinder added there, along with an insignificant subtraction for horsepower, and a huge dectractor of over 5.2 miles per gallon per added ton in weight, leaving merely an approximate **1.48 miles per gallon difference** for our choice in transmission, with a standard error nearly as large.

Isolating for transmission, number of cylinders, and overall vehicle tonnage, we see in *Figures 6, 7, and 8*, verification that the positive slope toward a manual transmission is dwarved by the negative slopes of the first and third most impactful regressors, being the vehicle's weight in tons and the number of cylinders invovled.

Given our probability value of ~**0.31**, having regressed for the other equivalent significant factors, this study **fails to reject the null hypothesis** that the choice of transmission makes any more than than a slight difference in fuel economy. Certainly nothing compared to the choice in vehicle weight.

Caveat

While this study does appear to have suggested a small but otherwise statistically insignificant increase to fuel efficiency between automatic and manual transmissions, this is not reliable information insofar as the data provides no single model of car with both transmission types. In the mid-70s, perhaps, a model had a transmission type, and not the choice in the matter that we have today, but because we do not know if automatic transmissions are not **heavier** than manual transmissions, it may well be that the possible weight lost in choosing a manual transmission might be completely explained by the regression in fuel economy given a vehicle's overall mass.

Source

Henderson and Velleman (1981), *Building multiple regression models interactively.* Biometrics, 37, 391–411.

Appendix

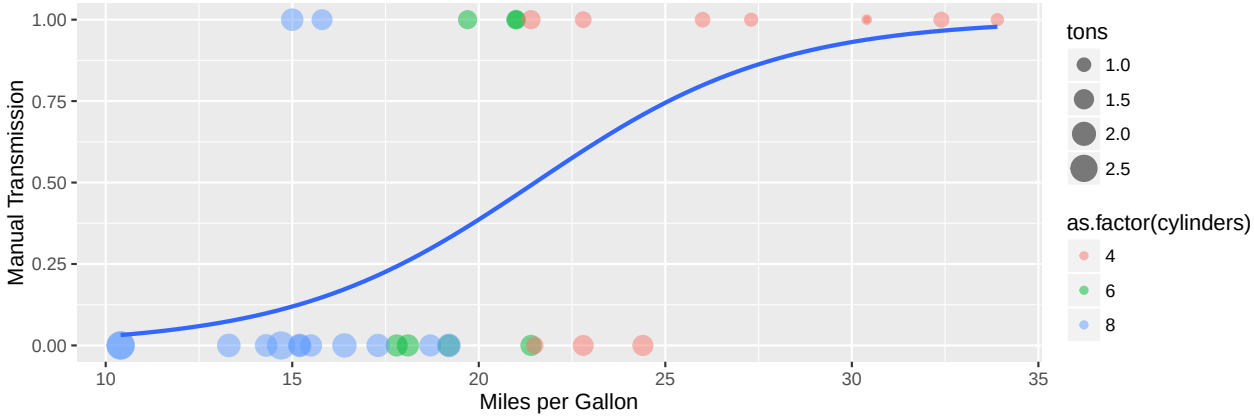
Figure 1 - Some of the data

	mpg	cyl	disp	hp	drat	wt	am	gear	carb
Mazda RX4	21	6	160	110	3.9	2.62	1	4	4
Mazda RX4 Wag	21	6	160	110	3.9	2.875	1	4	4
Datsun 710	22.8	4	108	93	3.85	2.32	1	4	1
Hornet 4 Drive	21.4	6	258	110	3.08	3.215	0	3	1
Hornet Sportabout	18.7	8	360	175	3.15	3.44	0	3	2
Valiant	18.1	6	225	105	2.76	3.46	0	3	1

Figure 2 - Summary of data after retitling

MPG	cylinders	horsepower	tons	manual
Min. :10.40	Min. :4.000	Min. : 52.0	Min. :0.7565	Min. :0.0000
1st Qu.:15.43	1st Qu.:4.000	1st Qu.: 96.5	1st Qu.:1.2906	1st Qu.:0.0000
Median :19.20	Median :6.000	Median :123.0	Median :1.6625	Median :0.0000
Mean :20.09	Mean :6.188	Mean :146.7	Mean :1.6086	Mean :0.4062
3rd Qu.:22.80	3rd Qu.:8.000	3rd Qu.:180.0	3rd Qu.:1.8050	3rd Qu.:1.0000
Max. :33.90	Max. :8.000	Max. :335.0	Max. :2.7120	Max. :1.0000

Figure 3 - A simple bionomial regression



Figures 4 - Simple Linear Model

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	17.15	1.125	15.25	1.134e-15
manual	7.245	1.764	4.106	0.000285

This regression suggests that your typical automatic will average about 17.15 miles per gallon, with your average manual transmission vehicle will clock it at over 24 miles per gallon. But…

Figure 5 - Multivariate Linear Model

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	27.76	2.723	10.2	9.364e-11
manual	1.478	1.441	1.026	0.3142
cylinders	-0.7452	0.5828	-1.279	0.2119
horsepower	-0.02495	0.01365	-1.828	0.07855
I(tons - mean(tons))	-5.213	1.84	-2.834	0.008603

Instead, shifting our intercept to that of the average weight vehicle, and regressive for cylinders, horsepower, and tonnage, we get a completely different picture, showing that after other factors are consider, very little detectable increase based on transmission alone remains.

Figures 6, 7, 8 - Individaulized Residual Plots