Regression Model Course - Course Project

Thiago Mata

29 January 2019

Executive Summary

The influence of the Tramission Types on the MPG is around0.5417548, from -1.1798384 to 2.263348 in the confidence interval of 95% . Despide of getting a positive value, when we consider our 95% confidence interval, there is a reasonable chance that the effect of the different types of Transmission Types is zero, positive or negative. In that sense, it **is inconclusive if some Transmission Type is better for MPG**. Using the T-Test over the model, model E, we can say, with p-value of 0.5253219, that knowning the Transmission Type changes the predictable result of the MPG, event after removing the influence of all others predictors.

Introduction

In this project, the goal is to make a study to the Motor Trend magazine that want know more about the relationship between a set of variables and miles per gallon (MPG) (outcome). The result should be presented as a PDF created from a Project R Markdown File They are particularly interested in the following two questions:

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

Analysis

As shown on the box plot, looking just to MPG and Type of Transmission seems to have a strong relationship between these two fields. So, different models are going to be created and compared to check this hypothesis.

With p-value of 0.8987 for Automatic Transmission Type and 0.5363 for Manual Transmission Type, we can assume that is a Normal Distribution and create linear models.

Creating Regression Models

Comparing Coeficients in Models

model	expectedOutcome	predictors
Model A	MPG	Just the Transmission Type
Model B	MPG	Fields strong correlated with the MPG, including Transmission Type
Model C	MPG	Fields strong correlated with the MPG, except Transmission Type
Model D	Transmission Type	Fields strong correlated with the Transmission Type, except MPG
Model E	Residuals of Model B	Just the Transmission Type

model	am	wt	vs	cyl	drat	hp	disp	RSS
Model A	7.2449	NA	NA	NA	NA	NA	NA	720.8966
Model B	1.9520	-3.2453	1.3976	-0.6399	0.5470	-0.0303	0.0135	158.6543
Model C	NA	-3.7231	0.5054	-0.9765	1.0451	-0.0238	0.0127	166.8170
Model D	NA	-0.2448	-0.4571	-0.1724	0.2552	0.0033	-0.0004	2.1422

As shown on the table above, some of the predictors of the Transmission Type (am), as Type of Engine (vs) and Weight (wt), are also predictors of the MPG. So, to have a clear vision about the influence of the Transmission Type on the MPG, let's see how good is the Transmission Type to predict the residual.

So, we can see that the Transmission Type, have a strong impact on the prediction. When we remove it, the Residual sum of squares - RSS increases 4.89%. But, the influence of it on the MPG is not as big in the model that only consider that field. So, many other fields also have a strong influence over the MPG, and some of them are also strongly correlated with the Transmission Type.

The Transmission Type has a strong influence on the MPG. In this model, we can say that only knowing that some car has Manual Transmission should increases our MPG prediction to 7.2449393 with residual squared error of 720.8965992. This high value is because, not only the Transmission Type may increase the MPG, but also because the Transmission Type has a strong coefficient of 0.255172 with the Weight (wt), -0.2447799, with the Rear Axle Ratio (drat), 0.255172,and with the Type of Engine (vs), -0.4570708. These fields also presents high coeficients to predict the MPG with -3.2453124, 0.5469634 and 1.3976089 respectivaly.

So, the Type of Transmission is a good to predict if a car presents many features that, all together, increase the car consume. Considering the 95% of confidence interval, that effects may vary in the interval of 3.6415096 to 10.848369.

Considering that we known everthing about a car. Then how much changing only the Transmission Type may affects the MPG? To answer this, we created a model that uses the most strong correlated fields. In this model, considering all the rest the same, changing a the Transmission type from Manual to Automatic should increase the MPG in 1.9520082 with residual squared error: 158.6543284. Considering the 95% of confidence interval, that effects may vary in the interval of -1.6735479 to 5.5775643.

After extracting all the predictability data from the others fields, this new model can measure if the remain data is still correlated to the Type of Transmission. The coefficient from the Type of Transmission in this case is 0.5417548. After removing the influence of the others fields, the influence of the Type of Transmission on the MPG is 0.5417548 with residual squared error of 164.5515427. Considering the 95% of confidence interval, that effects may vary in the interval of -1.1798384 to 2.263348.

As we can see on the residual plot, using just the Transmission Type to try to predict the MPG have a strong RSS and is not a good model. Adding the Transmission Type, on the model reduce the Residual Sum of Squareds in 4.89%.

Conclusion

If we don't have any other information about a car, beyond the Type of Transmission, then knowing that a car has Manual Transmission should have the MPG around 7.2449393 (from 3.6415096 to 10.848369 in the 95% confidence interval) bigger than the Automatic. If we know everything about a car, and just changing the transmission type from Automatic to Manual shoud increase the MPG around 1.9520082 (from -1.6735479 to 5.5775643 in the 95% confidence interval).

So, automatic cars in general consume more than manual ones. But, not necessarly because the Type of Engine increase the consume. But, in general, automatic cars presents others features, that also affects the consume. Considering the Transmission Type isolated, the real effect of changing the Transmission Type is 0.5417548 MPG (from -1.1798384 to 2.263348 in the 95% confidence interval).

Considering our 95% confidence interval, there is a reasonable chance that the effect of the different types of Transmission Types is zero. In that sense, it **is inconclusive if some Transmission Type is better for MPG**. But, we still can say that, if the Transmission Type affects the MPG, the impact is pretty low (from -1.1798384 to 2.263348).

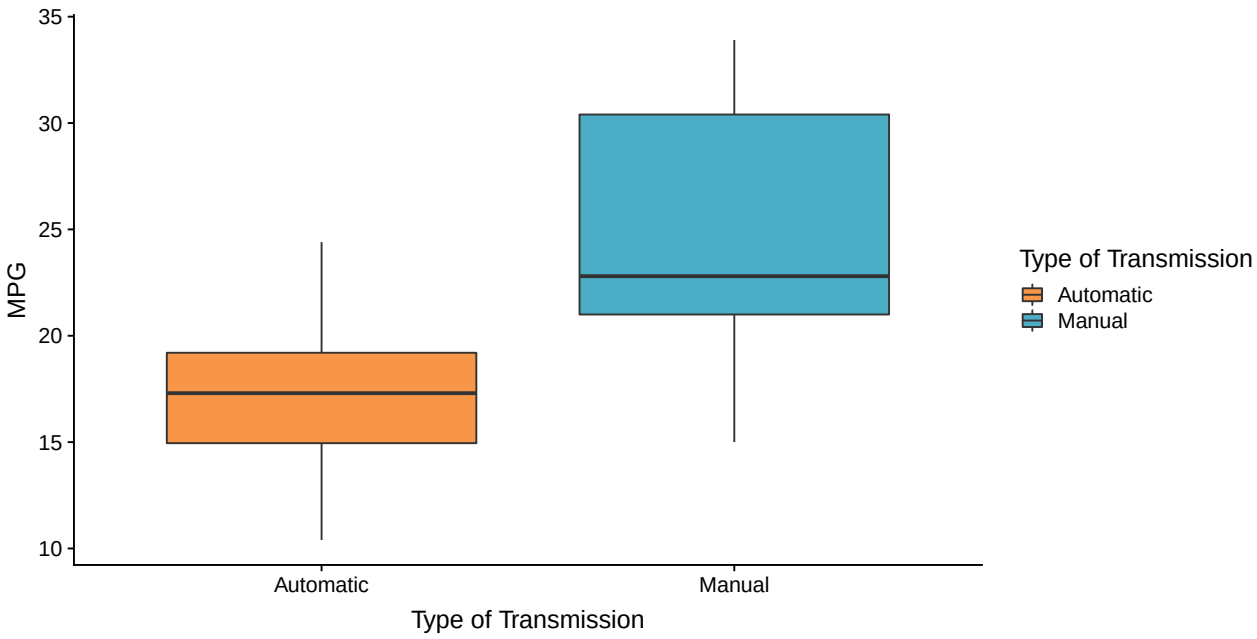
Appendices

Dataset

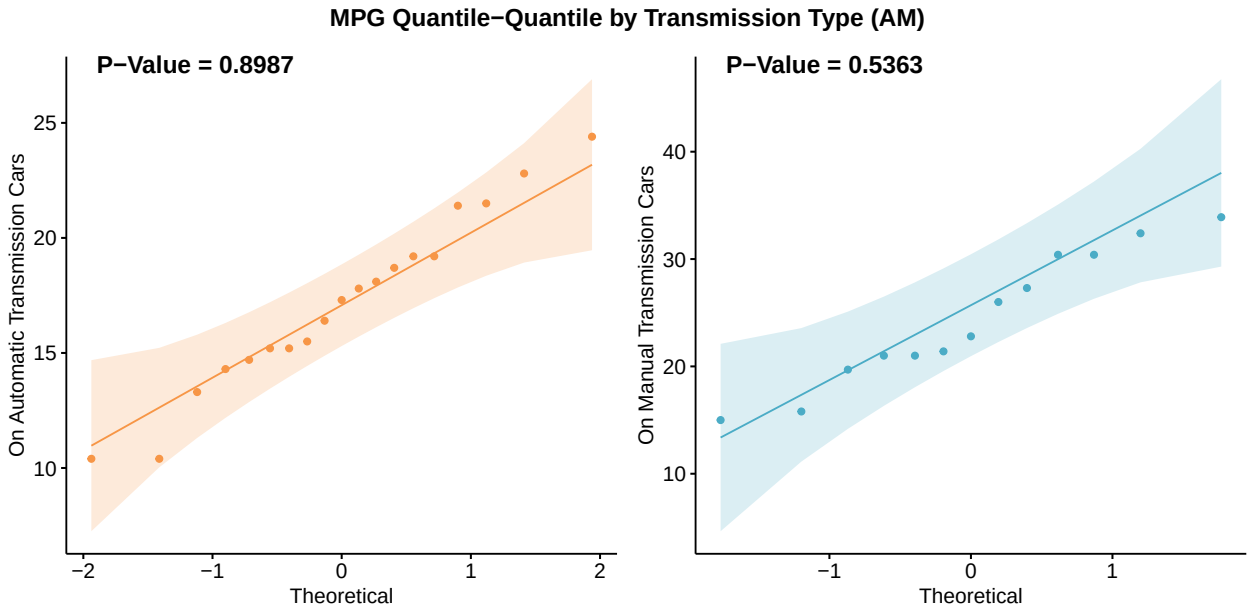
The dataset used in this study is the Motor Trend Car Road Tests Data. Each row contains the following fields:

field	description	percent.of.Null
mpg	Miles/(US) gallon	0 %
cyl	Number of cylinders	0 %
disp	Displacement (cu.in.)	0 %
hp	Gross horsepower	0 %
drat	Rear axle ratio	0 %
wt	Weight (1000 lbs)	0 %
qsec	1/4 mile time	0 %
vs	Type of Engine (0 = V-shaped 1 = straight)	0 %
am	Type of Transmission (0 = automatic 1 = manual)	0 %
gear	Number of forward gears	0 %
carb	Number of carburetors	0 %

Box Plot - Type of Transmission x MPG



Quantile-Quantile for Normality Check



Residuals of the Models