CENTRAL UNIVERSITY OF RAJASHTHAN, BANDARSINDRI

Introduction to Econometrics and Finance

Code – MBD 426

ASSIGNMENT REPORT

Regression Analysis on Breaking speed and stopping Distance of cars

Under Guidance of Dr. Pramod Kumar Naik

Submitted by:

Pradyumna Kumar Sahoo
2018MSBDA016
Department of Data Sciences and Analytics

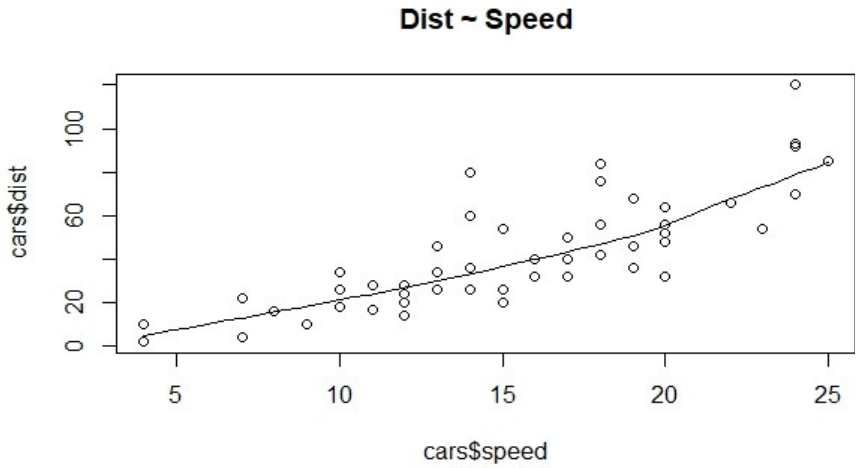
Data source

Inbuilt R data-set : Cars with 50 observations and 2 features
Y = Distance required to stop (in meter)
X = Speed of the car (in meter per second)

Exploratory data analysis

It is a good practice to analyze and understand the variables. Hence, I have shown first 5 observations from the data-set and also plotted the data-set.

Speed X(m)	Distance Y(ms^{-1})
4	2
4	10
7	4
7	22
8	16
9	10



Correlation

CORRELATION BETWEEN CAR SPEED AND STOPPING DISTANCE: **0.806**

Here as speed increases, the stopping distance also increases along with it, as there is a high positive correlation between them (closer to +1).

Building Linear Model

The aim is to establish a linear relationship (a mathematical formula) between the predictor variable (s) and the response variable, so that, we can use this formula to estimate the value of the response Y, when only the predictors (Xs) values are known .

$$\boldsymbol{distance = Intercept + (\beta * speed)}$$

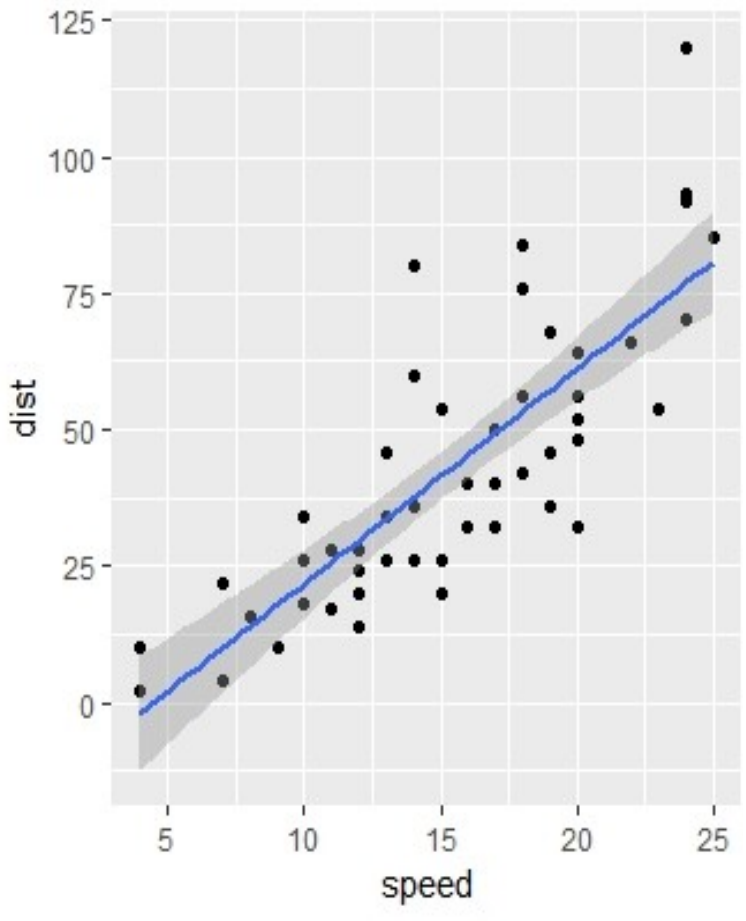
Model Summary – I

Variable	Coefficient	Standard Error	T - value	p-value	Level of Significance
intercept	**-17.57**	**6.75**	**-2.60**	**0.12**	*
speed	**3.93**	**0.41**	**9.58**	**$1.49 * 10^{-12}$**	***

You can notice the ‘Coefficients’ part having two components: **Intercept: -17.579, speed: 3.932**

These are also called the beta coefficients.

$$\textbf{distance} = -17.579 + 3.932 * \textbf{speed}$$



Graphical Representation of data with best fit line

Model Summary – II

Residual Standard Error	**15.38**
Degree of Freedom	**48**
R^2 Score	**0.65**
Adjusted R^2 Score	**0.64**
F statistic	**$89.57_{(1,48)}$**
Model P- Value	**$1.49*10^{-12}$**

Hypothesis Testing :

As per theory : The stopping distance is linearly related to the speed of the car.I,henceforth will try to determine whether the estimates satisfy the expectations of the theory being tested or not. For that,I will use a significance level of 5%.

My Null Hypothesis (H_0) and Alternate Hypothesis(H_1) are as follows:

H_0 = Stopping distance does not depend on car's speed **(β =0)**

H_1 = Stopping distance does depend on car's speed. **(β ≠0)**

In order to prove my hypothesis, **T- test** is employed:

$$t = \frac{\hat{\beta} - \beta^*}{S.E(\hat{\beta})}$$

Where $\hat{\beta}$ is actual β^* we have calculated from the model and β^* is value from null hypothesis and S.E($\hat{\beta}$) is standard error of $\hat{\beta}$.In this case we are calculating for $\hat{\beta}$ and from the summary,we have the value of $\hat{\beta}$ and S.E($\hat{\beta}$) and from the null hypothesis we have $\beta^* = 0$. Therefore,

$t_{(1,48)}$= (3.93-0)/0.41 =9.58

Here the level of significance is 0.05 and from above t-Test we have calculated t=**9.58** with

degree of freedom 48 which is greater than t-critical=2.0.

9.58 > 2.0

This implies that we can ***reject the Null Hypothesis*** i.e. Stopping Distance is dependent on Car's speed.

Inference drawn:

- There exists a linear relationship between stopping distance and car's speed just like theory predicted.
- Stopping Distance is dependent on Car's speed.

THANK YOU