Experiment No 2: Program on Logistic Regression

Abstract

It measures the relationship between the categorical dependent variable and one or more independent variables by estimating probabilities using a logistic function and has values of the response variable as in 0 or 1, True or False etc based on the mathematical equation referencing predictor variable. The interpretation of logistic regression can be done using numerical analysis.

Introduction

The objective of this program is to conclude on the probability of manual transmission of vehicles and the general mathematical equation for logistic regression is as follows:

y = 1/(1+e^-(a+b1x1+b2x2+b3x3+...)

Here y is the response variable and x is the predictor variable. a and b are the coefficients which are numeric constants. The function that is used to create the regression model is the glm () function i.e.glm- means generalized linear model which constitutes of formula, data and family.

The in-built data set "**mtcars**" used as for this experiment describes different models of a car with their various engine specifications. In "mtcars" data set, the transmission mode has a binary value of (0 or 1). We can create a logistic regression model between the columns "am" and 3 other columns - hp, wt and cyl.

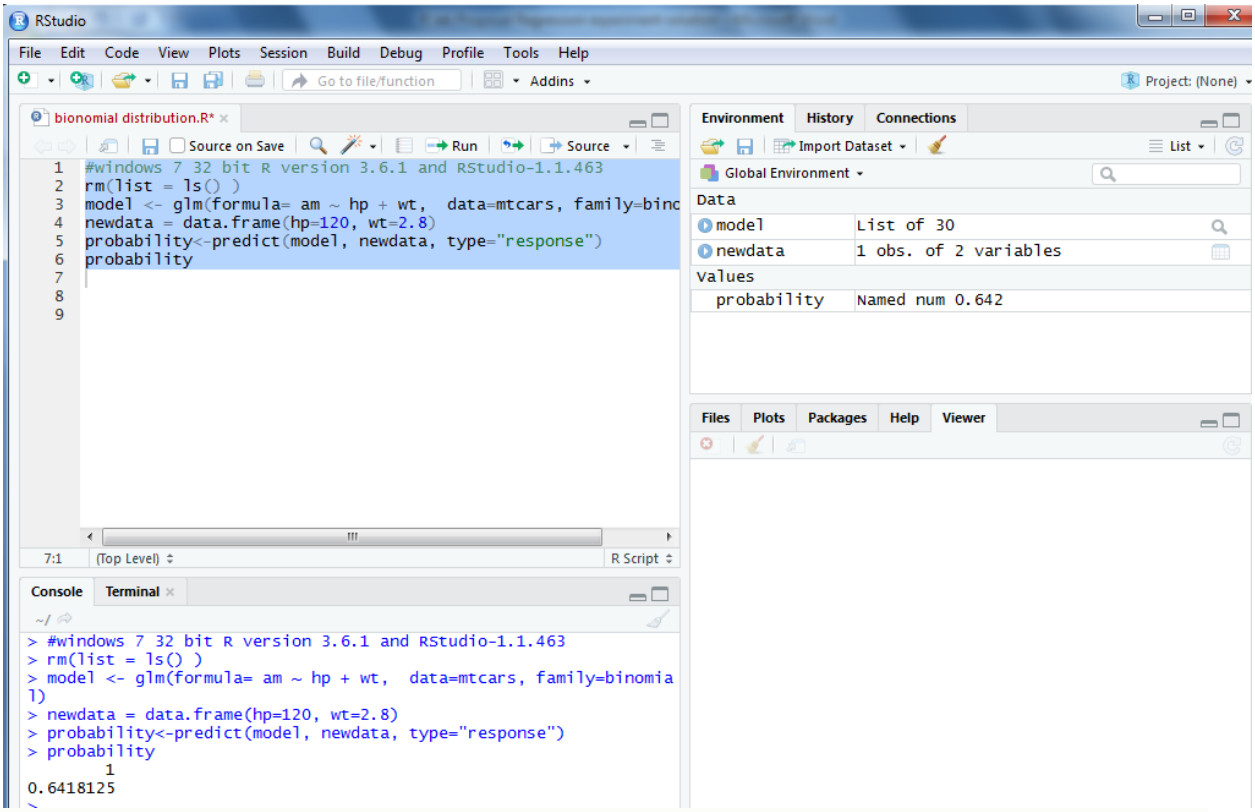
Experiment

This experiment focuses on dataset of ‘mtcars’ to find the probability of vehicle fitted with manual transmission.By use of the logistic regression equation of vehicle transmission in the data set mtcars, estimate the probability of a vehicle being fitted with a manual transmission if it has a 120hp engine and weights 2800 lbs.

Solution Steps

1) Numerical Analysis with R

We apply the function glm to a formula that describes the transmission type (am) by the horsepower (hp) and weight (wt). This creates a generalized linear model (GLM) in the binomial family.



Conclusion: For an automobile with 120hp engine and 2800 lbs weight, the probability of it being fitted with a manual transmission is about 64%.