Introduction to Computational Statistics INSH 5301

Homework 07

02/24/2020

PLEASE copy and paste the whole question number and text into submission so I can grade easily. When I grade easier, you might get better grade!

Load the mtcars dataset in R.(use data(mtcars)) Using this sample of cars, we are interested in knowing if:
(a) there is any relation between the miles per gallon (mpg) and the gross horsepower (hp),
(b) the sign of the relation, and,
(c) the magnitude - strenght - in which this variables are associated.

1. Using ggplot2, make a scatter plot with mpg on x-axis and hp on the y-axis. Using the graph as evidence, do you think the variables are positively, negatively or not related?

2. Compute the covariance between the two variables. How can we use the covariance to answer (a,b,c)?

3. Calculate the correlation between the two variables. How can we use the correlation coefficient to answer (a,b,c)?

4. Estimate the regression coefficients β_0 and β_1. What's the interpretation of β_1?

5. Write a function that, using the estimated coefficients, returns the fitted value of hp given some mpg. Find all the fitted values using all observations of mpg. Make a scatterplot of mpg and the fitted values.

6.Calculate the errors of the regression, i.e. compute the difference between the fitted values and the observed values. Make an histogram with the errors. Do they look normally distributed?

7.Compute the standard error of β 1 and use a t-test to check wheter β 1 is significant - statistically different than zero -. Calculate the p-value and 95% confidence interval to verify that you reach the same conclusion.

8. Calculte the R^2 coefficient and interpret it's value

9. Make a new graph with ggplot2, this time make a scatter-plot with the regression line and a 95% C.I. for the fitted values.