26 BE 7023 & 26 PH 7023

Advanced Biostatistics

Fall 2018

MB Rao

Homework Sheet No. 2 Due Date: September 19, 2018 Maximum Points: 30

Download the data ‘Animals’ from the ‘MASS’ package. This data is special. The first column identifies the animals involved in the study. This column is not a variable in the traditional sense. It is a column of labels.

1. What is the dimension of the data? 2 points

2. Describe the data. 2 points

3. Show the scatter plot of the data with x-axis ‘body’ and y-axis ‘brain.’ Identify as many data points as you can. Comment on the plot. 4 points

4. Show the scatter plot of the data after the logarithmic transformation. Identify as many points as you can. Comment on the plot. Obtain the simple linear regression model. Draw the line on the scatter plot. Make the graph as informative as possible. 7 points

5. Remove the dinosaurs. (Do you know how to do it?) Show the scatter plot of the resultant data after the logarithmic transformation. Identify as many points as you can. Comment on the plot. Obtain the simple linear regression model. Draw the line on the scatter plot. Make the graph as informative as possible. 7 points

6. Write the prediction model coming from Question 5 directly in terms of the original variables. 2 points

7. Calculate the ratio of the model from Question 6 for all animals in the data. Arrange the ratios in increasing order of magnitude. Comment on the output.

4 points

8. Comment on the graphs in Questions 5 and 6. 2 points