26 BE 7023 & 26 PH 7023: Advanced Biostatistics

Autumn 2018

MB Rao

Homework Sheet No. 1 Due date: September 12, 2018 Maximum points: 30

Simple linear regression + confidence and prediction bands

Download and activate the package ‘faraway.’ Download the data ‘prostate.’

1. Describe the data (?prostate on R). Include the size, top ten rows and summary statistics of the data. 4 points
2. Obtain a plot of the data with x = lcavol and y = lpsa. Comment on the graph. 4 points
3. Fit a simple linear regression model by taking Y = lpsa and X = lcavol. Write the prediction equation. Report R2and comment on it. Estimate the population standard deviation. 4 points
4. Write a short note on prostate specific antigen (PSA). 2 points
5. Transform the regression model back to the original variables. Comment on the resultant model. 4 points
6. Build a 95% confidence band as well as prediction band around the regression line. Make the graph as descriptive as possible. 5 + 5 + 2 points