STAT 415/615 Classwork/Lab Week 4 Session 1

Residuals, Residual Plots, Transformations.

Instructions: Submit to Canvas an R Markdown file and a Word file that show all script, R code, and requested output by the due date for this assignment.

1) Page 152 problem 3.24 a, b

2)

A picture containing table

Description automatically generated

3a) For the bivariate data given below, use r coding to produce a scatter plot and a residual plot (residuals against x) Examine the two plots and confirm that a linear model is not appropriate.

3b) Perform a square root transformation on y , plot your new data and discuss your now plot. Is it closer to being more appropriate for a linear regression model ?

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **x** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| **y** | 2 | 1 | 6 | 14 | 15 | 30 | 40 | 74 | 75 |
|  |  |  |  |  |  |  |  |  |  |

4)

For the bivariate data given below, the residual plots suggest problems involving non-normality or non-constant error variance or both. Use Box – Cox method as indicated in class to produce a lamda power transformation that will best normalize the data. Your work should include all graphs and plots that support your final answer. Use a series of steps and reasoning demonstrated in class.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| x | 7 | 7 | 8 | 3 | 2 | 4 | 4 | 6 | 6 | 7 | 5 | 3 | 3 | 5 | 8 |
| y | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 3 | 3 | 3 | 6 | 7 | 8 |