Chapter 3 Question 16

Written Out Answers

1. See the knitted section for the scatter plot of the data. According to Figure 3.15, the best transformation on Y would be Y’ = log10 Y in order to achieve constant variance and linearity.
2. From the Box-Cox plot in the knitted section, it is clear that the best choice for λ is 0 which corresponds to Y’ = log10 Y.
3. Using the transformation Y’ = log10 Y, the estimated linear regression function for the transformed data is Ŷ ‘ = -0.1954003 X + 0.6548798.
4. See the plot of the estimated regression line and the transformed data. The regression line does seem to be a good linear fit of the transformed data.
5. The residuals for the transformed data are:

|  |  |
| --- | --- |
| 1 | -0.051178946 |
| 2 | 0.057965523 |
| 3 | 0.006813001 |
| 4 | -0.082957620 |
| 5 | -0.056628681 |
| 6 | 0.035141692 |
| 7 | 0.012317861 |
| 8 | 0.085549775 |
| 9 | 0.046397651 |
| 10 | 0.017680995 |
| 11 | -0.007980995 |
| 12 | -0.039295058 |
| 13 | -0.006161112 |
| 14 | -0.049546328 |
| 15 | 0.031882242 |

See the knitted section for a plot of the residuals

against the transformed fitted values. Also see the  
  
knitted section to see the normal probability plot.

The residuals against fitted values plot shows the variance is much more constant and much closer to 0 which means the transformation was successful in achieving constant variance. The normal probability plot shows a much more linear plot which shows the transformation was successful in achieving linearity.

1. The estimated regression function in the   
   original units is log10 (Ŷ) = -0.1954003 X + 0.6548798.