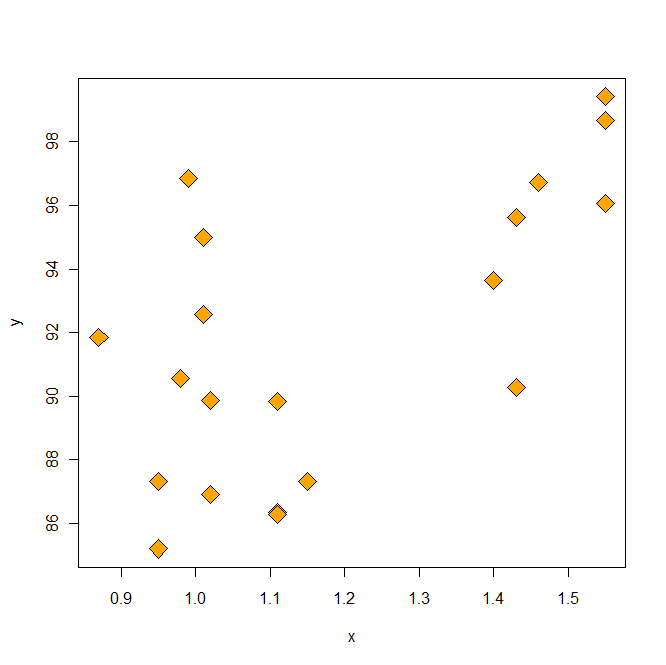
linear model – chapter 2 excerciseS

**Thanh Doan – Student ID 0159701**

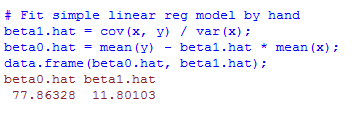
|  |  |
| --- | --- |
| exercise 2.7 |  |

**a - Fit a simple linear model to the data**

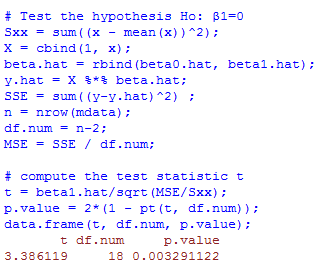
* It might be useful to plot a scatter diagram to visualize the data first.



* Fit the linear model. Compute least-square estimation of the parameters by hand (i.e. not using **lm(.)** function provided by R statistical computing language).

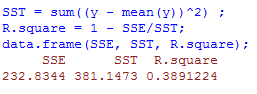


**b – Test the hypothesis**

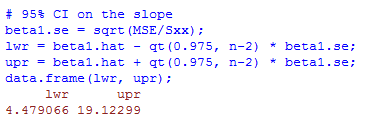


P-value = 0.003. The null hypothesis is rejected at the confidence level 95% (or even 99%) and we conclude that there might be a linear relationship between purity and hydrocarbons.

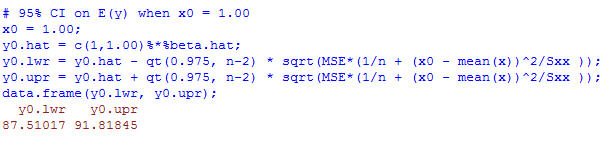
**c – Calculate**



**d – Find 95% CI on the slope**

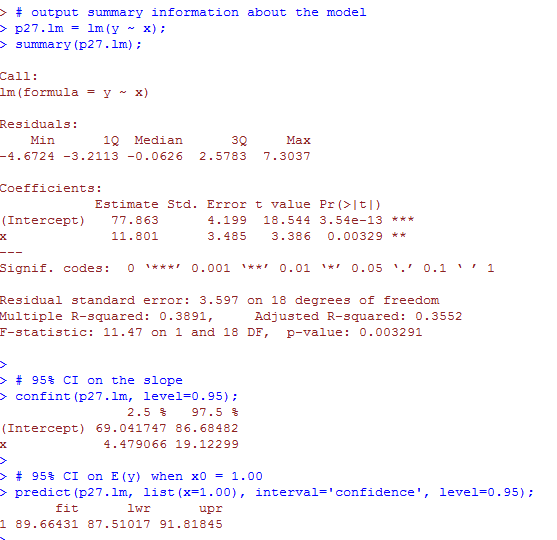


**e – Find 95% CI on the mean y when x = 1.00**



**NOTE**

In practice, we can answer all the 5 questions above easier by using much fewer lines of R code. I.e. if we use convenient built-in functions in R to analyze this dataset, then we can accomplish all the above work with just 4 lines of R code below.



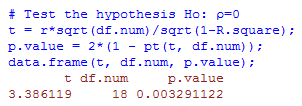
|  |  |
| --- | --- |
| exercise 2.8 |  |

**a – Find the correlation between x and y**

* The correlation

**b – Test the hypothesis**

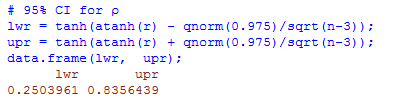
* This test is basically the same test



Thus, and we conclude that there might be a linear relationship between purity and hydrocarbons.

**C – Find 95% CI for**

The 95% confidence interval for is

****

|  |  |
| --- | --- |
| exercise 2.19 |  |

**a –**