University of California, Los Angeles
Department of Statistics

Statistics 13 **Instructor: Nicolas Christou**

Lab 6

Exercise 1:
Use `R` to access the Maas river data. These data contain the concentration of lead and zinc in `ppm` at 155 locations at the banks of the Maas river in the Netherlands. You can read the data in `R` as follows:

```
a <- read.table("http://www.stat.ucla.edu/~nchristo/statistics13/soil.txt", header=TRUE)
```

Answer the following questions:

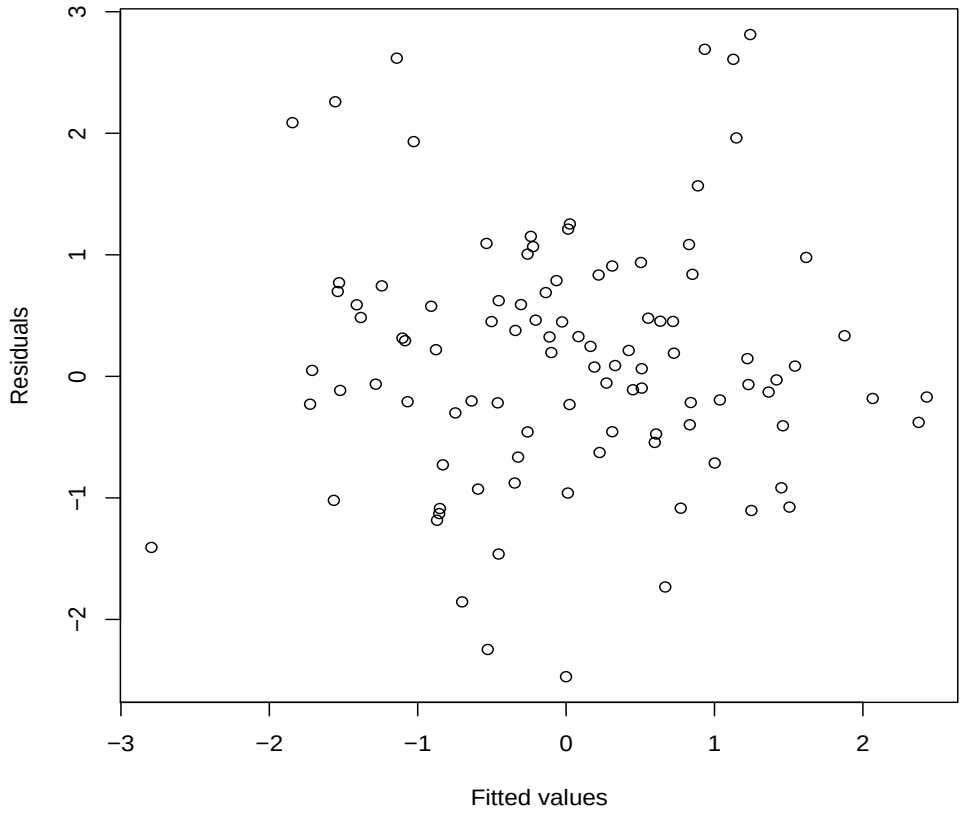
a. Run the regression of `lead` on `zinc`. Submit the `R` printout. What is the value of R^2.

b. Run the regression of `log(lead)` on `log(zinc)`. Submit the `R` printout. What is the value of R^2.

c. Test the hypothesis that the slope of the model in question (b) is equal zero against the alternative that it is not equal zero. Use the p-value from the `R` output.

d. Repeat question (c): Do all the calculations by hand, i.e. compute the t statistic using the formula from your class notes, and compare it to a critical t value with $\alpha = 0.05$.

e. Plot the residuals (`y`-axis) against the fitted values (`x`-axis) of the model of part (b). Are there any violations of the assumptions of regression? In general this plot must look like a patternless cloud of points (see an example below).



Exercise 2:
The data

```
a <- read.table("http://www.stat.ucla.edu/~nchristo/statistics13/chestnut.txt", sep=",", header=TRUE)
    x   y
1   4 0.8
2   5 0.8
3   8 1.0
4   8 3.0
5  10 2.0
6  10 3.5
7  12 4.9
8  13 3.5
9  14 2.5
10 16 4.5
11 18 4.6
12 20 5.5
13 22 5.8
14 23 4.7
15 25 6.5
16 28 6.0
17 29 4.5
18 30 6.0
19 30 7.0
20 33 8.0
21 34 6.5
22 35 7.0
23 38 5.0
24 38 7.0
25 40 7.5
26 42 7.5
```

give the diameter at breast height (y) and the the age (x) of 26 chestnut trees (from *Mathematical Statistics and Data Analysis* by John Rice, 3rd Edition). Answer the following questions:

a. Try fitting the diameter at breast height as a linear function of age: $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$. What percent of the variation in y can be explained by x?

b. Refer to question (a): Test the hypothesis that $\beta_1 = 0$ against the alternative that $\beta_1 \neq 0$. Use $\alpha = 0.05$.

c. Refer to question (a): Now examine the residuals. What do you observe? Can you suggest a transformation that will produce a more linear relationship?

d. Fit the diameter at breast height as a linear function of the square root of age: $y_i = \beta_0 + \beta_1 \sqrt{x_i} + \epsilon_i$. What percent of the variation in y can be explained by the predictor variable?

e. Refer to question (d): Examine the residuals. What do you observe?

f. Refer to question (d): Test by hand using the formula fro your class notes the following hypothesis:
$H_0 : \beta_1 = 1.3$
$H_a : \beta_1 \neq 1.3$

g. Test the hypothesis of question (f) using the permutation test (see homework 2 - exercise 8).