Statistical Modeling I
Practical in R

Practical in R

This practical reminds you how to load the data and how to run a linear regression in R. Moreover, it gives you the opportunity to learn how to work with standardized residuals and fitted values.

To calibrate an instrument which measures the lactic acid concentration in blood, an investigator used twenty samples of known concentration (x) and then obtained the concentration as measured on the instrument (y). The results were

x	1	1	1	1	3	3	3	3
y	1.1	0.7	1.8	0.4	3.0	2.4	4.9	4.4
x	5	5	5	5	10	10	10	10
y	4.5	7.3	8.2	6.2	12.0	13.1	12.6	13.2
x	15	15	15	15				
y	18.7	19.7	17.4	17.1				

Enter the x and y values as vectors in R.

1. Produce a scatterplot of the data. Does the relationship between y and x seem to be linear?

2. Fit the model to the data by `conc <- lm(y ~ x)`
   we use `conc` as the name of our model as the data is about concentration of lactic acid.

   To see the details of our fitted model we use
   `summary(conc)`

3. Add the fitted line to your scatterplot by the following command
   `plot(x,y, main="Fitted Line Plot")`
   `(abline(conc))`

4. See the Analysis of Variance table by using the following command
   `anova(conc)`

5. We need to store the standardised residuals for plots
   `stdres <- rstandard(conc)`
   `print(stdres)`
   `plot(x,stdres, main="Std residuals versus x")`
   and save this plot.

   Is there any reason to doubt the linearity of the model?

6. We also save the fitted values
```
fits <- fitted(conc)
print(fits)
plot(fits,stdres, main="Std residuals versus fits")
```
save this plot.

Is there any reason to doubt that the variance is constant?

7. Finally we look at a Q-Q plot
```
qqnorm(stdres, main="Q-Q Plot")
qqline(stdres)
```
save this plot.

Is there any reason to doubt the assumption that the errors are normally distributed?