Statistical models

(STATS240.1: A short course in R)

Gunnar Stefánsson, Ásta Jenný Sigurðardóttir and Lorna Taylor

October 15, 2012

Linear statistical models

Mathematical model:

$$y = \alpha + \beta x + \epsilon$$

R definition:

$$y \sim x$$

`lm(y~x)`

Storing the output

`fm<-lm(y~x)` .

A sequence:

`fm<-lm(y~x)`	`Fitting the model`
`summary(fm)`	`Traditional summary`

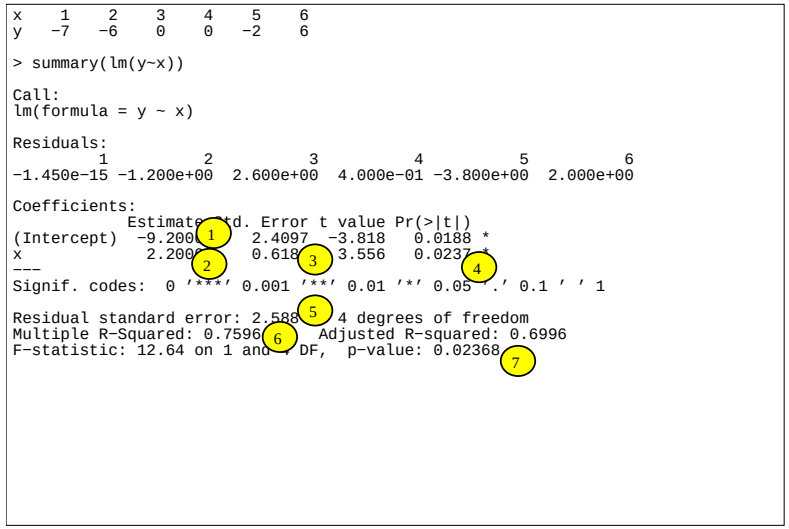

Figure: Example output from a simple linear model fit of the form y=a+bx. Items (1)-(2) are the estimates of a and b respectively. The estimate of the standard error of b is given by (3). The P-value for testing whether the true (underlying) value of b is zero is in (4). Items (5)-(7) give the MSE, R-squared and P-value for the entire model, respectively.

Nonlinear statistical models

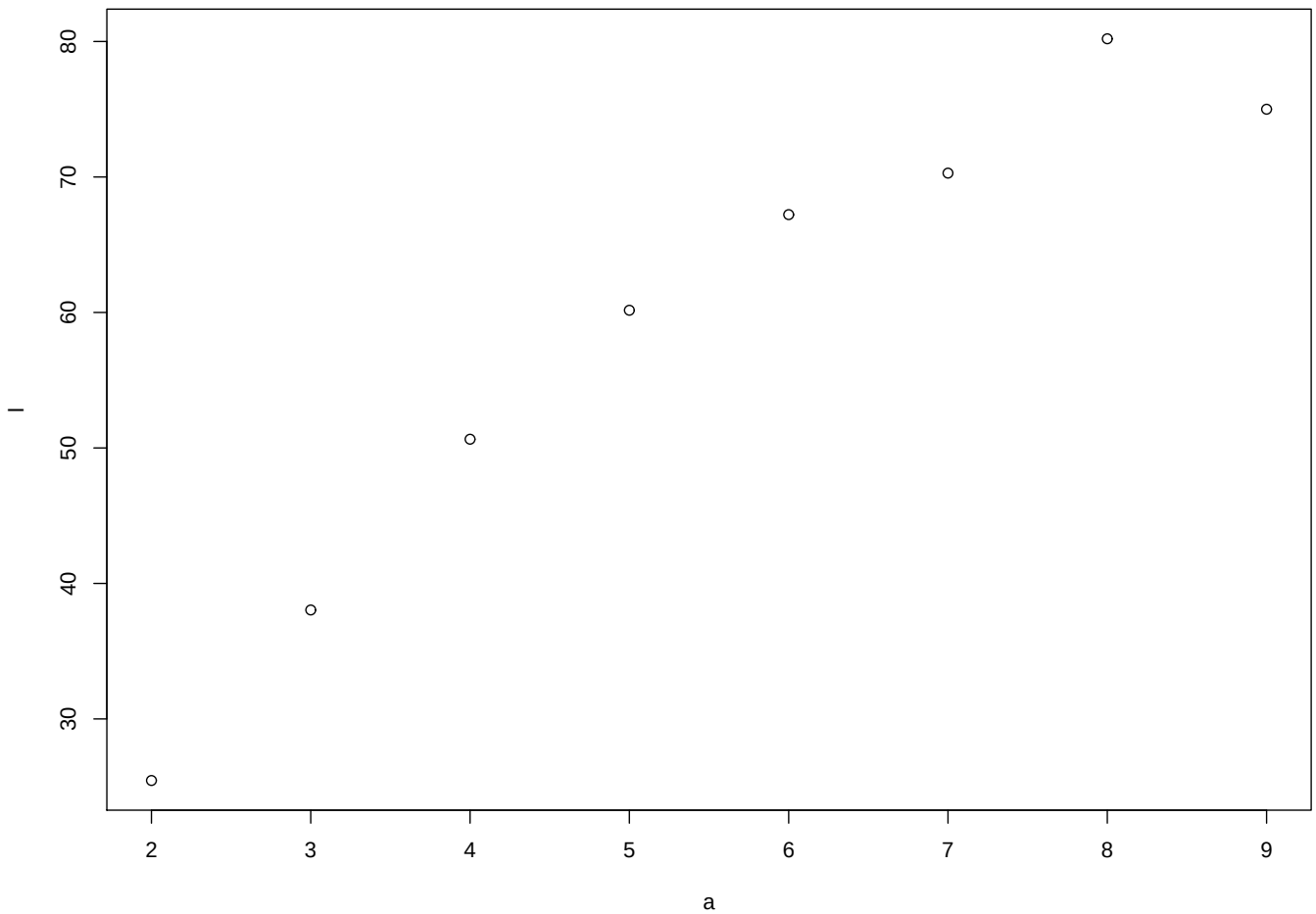

Figure: Example of a potential nonlinear relationship (length and age).

Miscellaneous statistical models

Usual assumptions: Linear, Gaussian errors, constant variance, independence.
Alternatively: Nonlinear, non-Gaussian, heterogeneity, non-independence.
Examples: Length-weight relationships, spatial correlations etc

Further reading

Extensive references exist for statistical models in R (or Splus).
For simple statistical analyses the statistical basis can be obtained from any introductory book (such as Moore and McCabe's).
For linear models in R, consult any corresponding textbook, (such as Fox et al).
For more detailed regression analysis a comprehensive book on linear models (such as John Neter et al) is needed.