Computer Lab Session, Teaching Week 8

Aims of this lab session:

• to familiarise yourself with R’s methods for fitting non-linear regression models.

All data sets used in this lab are available in the Data3S6 package (version 3.4 or later). Thus, to access these data sets, you should issue the command

library(Data3S6)

at the start of your R session.

1 Puromycin data

First read the help file for the Puromycin data set. Next, extract the subset of data for the untreated cells only:

dat <- subset(Puromycin, state=="untreated")

We will now fit the Michaelis–Menten model

\[ Y_i = \frac{\beta_0 x_i}{\beta_1 + x_i} + \varepsilon_i, \quad i = 1, \ldots, n \]

to these data, where rate is taken as the response variable \( Y \) and conc is the regressor variable \( x \).

Fit the Michaelis–Menten model using each of the following three methods:

(A) Use the function \texttt{nls()} directly, i.e. code the model as a non-linear model formula and use \texttt{nls()} with a suitable vector of starting values for the parameters of the model;

(B) use \texttt{nls()} as above, but add the argument \texttt{algorithm="plinear"}, which uses a different algorithm to minimise the RSS, exploiting the fact that the model is linear with respect to one of its parameters; and

(C) use the “self-starting non-linear model” \texttt{SSmicmen} for the Michaelis-Menten equation. A self-starting model is an R object, essentially a model equation equipped with a rule for computing starting values. To apply this, call

\texttt{nls(rate ~ SSmicmen(conc, b0, b1), dat)}

without a vector of starting values.

Assessment Questions:

1. What are the fitted parameter values for \( \beta_0, \beta_1 \) and \( \beta_2 \) using method A?
2. Are the fitted models, obtained using the three methods listed above, identical or different?

3. Plot the data as a scatterplot, and using predict, add the fitted curve obtained using method C.

If you are present at the lab class, please show the plot to one of the lab demonstrators who will check that it is correct.

3. Plot the data as a scatterplot, and using predict, add the fitted curve obtained using method C.

If you are present at the lab class, please show the plot to one of the lab demonstrators who will check that it is correct.

If you are absent from the lab class, please attach a hard copy printout of the plot.

2 Nonlinear Regression for Available Chlorine Data

A study conducted by Procter & Gamble reported the amount of available chlorine in a cleaning product as a function of time since manufacture. After loading Data3S6 you can access these data with the command

```r
chlorine
```

The data frame chlorine contains variables Weeks (the time in weeks since the product was manufactured) and Chlorine (the amount of available chlorine). Produce a scatter-plot of the data.

If we denote the amount of chlorine at the ith observation by $Y_i$ and the time by $t_i$, then a postulated theoretical model for the relationship between the variables is

$$Y_i = \beta_0 + \beta_1 e^{-\beta_2 t_i} + \varepsilon_i$$

where $\varepsilon_1, \ldots, \varepsilon_n$ are independent $N(0, \sigma^2)$ random variables.

The aim of this exercise is to estimate the model parameters for the chlorine data using non-linear regression.

Assessment Questions:

4. Plot the data.

If you are present at the lab class, please show the plot to one of the lab demonstrators who will check that it is correct.

If you are present at the lab class, please show the plot to one of the lab demonstrators who will check that it is correct.

If you are absent from the lab class, please attach a hard copy printout of the plot.

5. To apply nls we will need to choose starting values for the parameters. We can get an initial value for $\beta_0$ by noting that this parameter represents the asymptotic minimum amount of chlorine available. Inspect the plot, and read off an approximate value for $\beta_0$.

If you are absent from the lab class, please attach a hard copy printout of the plot.
6. Having got the initial value of $\beta_0$, say $b_0$, note that
\[ \log(Y_i - b_0) \approx \log(\beta_1) - \beta_2 t_i. \]
With a bit of thought you should now be able to find suitable initial values for $\beta_1$ and $\beta_2$.

7. Now apply \texttt{nls} and fit the model. Report the fitted parameters:

8. Plot the fitted curve and the original data.
   If you are present at the lab class, please show the plot to one of the lab demonstrators who will check that it is correct.

9. Visually assess the quality of the fitted model.

10. According to the fitted model, how long does it take for the product to lose 10% of the chlorine that was initially available?
11. Is this model partially linear? (i.e. can the “plinear” algorithm be used for this problem?) Explain why.

12. Find out whether the R package includes a self-starting non-linear model that is similar to the one we fitted here.