

Introduction to Statistical Practice

Assessment, 1999–2000

The data in the following Table, collected by Brian Everitt, are described in the Handbook of Small Data Sets (HSDS) as being the ‘weights, in kg, of young girls receiving three different treatments for anorexia over a fixed period of time with the control group receiving the standard treatment’.

| Cognitive behavioural treatment | | Control | | | | Family therapy | |
|---------------------------------------|-------|-------------|-------|-------------|-------|-------------------|-------|
| Weight (kg) | | Weight (kg) | | Weight (kg) | | Weight (kg) | |
| before | after | before | after | before | after | before | after |
| 80.5 | 82.2 | 80.7 | 80.2 | 83.8 | 95.2 | | |
| 84.9 | 85.6 | 89.4 | 80.1 | 83.3 | 94.3 | | |
| 81.5 | 81.4 | 91.8 | 86.4 | 86.0 | 91.5 | | |
| 82.6 | 81.9 | 74.0 | 86.3 | 82.5 | 91.9 | | |
| 79.9 | 76.4 | 78.1 | 76.1 | 86.7 | 100.3 | | |
| 88.7 | 103.6 | 88.3 | 78.1 | 79.6 | 76.7 | | |
| 94.9 | 98.4 | 87.3 | 75.1 | 76.9 | 76.8 | | |
| 76.3 | 93.4 | 75.1 | 86.7 | 94.2 | 101.6 | | |
| 81.0 | 73.4 | 80.6 | 73.5 | 73.4 | 94.9 | | |
| 80.5 | 82.1 | 78.4 | 84.6 | 80.5 | 75.2 | | |
| 85.0 | 96.7 | 77.6 | 77.4 | 81.6 | 77.8 | | |
| 89.2 | 95.3 | 88.7 | 79.5 | 82.1 | 95.5 | | |
| 81.3 | 82.4 | 81.3 | 89.6 | 77.6 | 90.7 | | |
| 81.3 | 82.4 | 81.3 | 89.6 | 77.6 | 90.7 | | |
| 76.5 | 72.5 | 78.1 | 81.4 | 83.5 | 92.5 | | |
| 70.0 | 90.9 | 70.5 | 81.8 | 89.9 | 93.8 | | |
| 80.4 | 71.3 | 77.3 | 77.3 | 86.0 | 91.7 | | |
| 83.3 | 85.4 | 85.2 | 84.2 | 87.3 | 98.0 | | |
| 83.0 | 81.6 | 86.0 | 75.4 | | | | |
| 87.7 | 89.1 | 84.1 | 79.5 | | | | |
| 84.2 | 83.9 | 79.7 | 73.0 | | | | |
| 86.4 | 82.7 | 85.5 | 88.3 | | | | |
| 76.5 | 75.7 | 84.4 | 84.7 | | | | |
| 80.2 | 82.6 | 79.6 | 81.4 | | | | |
| 87.8 | 100.4 | 77.5 | 81.2 | | | | |
| 83.3 | 85.2 | 72.3 | 88.2 | | | | |
| 79.7 | 83.6 | 89.0 | 78.8 | | | | |
| 84.5 | 84.6 | | | | | | |
| 80.8 | 96.2 | | | | | | |
| 87.4 | 86.7 | | | | | | |

Table 1: Anorexia data

Produce a report on the above data-set, and hand it in to me by **Friday of First Week, next Term**.

The main problem is to compare the methods of treatment. However, you may think of other important questions and decide to investigate them also.

You can use the S-Plus Help to learn more about **fitting models**. In S-Plus 2000 (or S-Plus 4.5), click successively on: ‘Help’, ‘Splus Help’ and then the ‘Contents’ section. This will take you to a list including

information on for example ‘Fitting Statistical Models’ (including linear regression etc.), ‘Building Formulas’ and ‘Working with Script Commands and Report Windows’.

You may find ‘script commands’ useful—they represent the language used by older versions of S, as in Robin Reed’s handout. You can use them to communicate directly with another application, such as a database, spreadsheet or programming environment. Most importantly for your purposes, you can save scripts to disk for later use, and will usually be able to run them in older versions of S-Plus (perhaps after minor modifications). You can search for details of the commands under ‘Help’, ‘Search’. For example, look for ‘~’ and ‘formula’.

In your report, include a full **discussion of the data and of your analysis**. For example,

- Are there any unusual features in the data? Are the three groups similar before treatment? How reliable are the data?
- Include appropriate plots of the data. For example, HSDS says ‘it is instructive to look at the three scatterplots of after/before’.
- Do you think it appropriate to transform the data, or modify the data-set in any other way (for example, by omitting any outliers)? Why or why not?
- What do you consider to be a suitable measure of success of treatment, and why?
- Give a simple model for predicting the effects of each of the three treatments. Under what circumstances do you believe your model would work well, for example would you use the same model over the whole range of ‘before’ weights?
- What other information and explanatory variables do you think might help in choosing the best treatment?

Use ‘ \LaTeX ’ (preferably!), or some other package like ‘Word’ to produce your report, and demonstrate that you can use various **different features of S** [try to show off here!]

There are no lower or upper limits to the length of your report; stop when you feel you have done as much as you usefully can in interpreting the data, and perhaps say why you stopped where you have.

Good Luck!