University of York

Department of Health Sciences

Applied Biostatistics

Sample Assessment

This is the assessment which was used last year. Each student was given a slightly different version of the data, so each student's submission should be different. You can download one sample set from the website.

A sample answer will be posted later in the term, when we have covered enough statistics to do this assignment.

Applied Biostatistics Assignment 2005/2006

Psoriasis is a chronic skin disease that can affect people of all ages. One method of measuring it is to use the Psoriasis Area and Severity Index (PASI).

A trial was carried out to examine the effect of a new form of treatment for psoriasis. The following variables were measured and are contained in the SPSS datafile.

Sex: Sex of the patient (0 = female, 1 = male).

Age: Age (in months)

Years: Number of years since the person was first diagnosed with psoriasis.

PASI: Psoriasis Area and Severity Index – a measure of disease severity, taking into account both the amount of coverage and the severity of coverage. The maximum score on the PASI is 200.

Group: Treatment or control group (1 = treatment)

Postpasi: The PASI measured 6 months after treatment.

Improved: The person's subjective opinion as to whether their psoriasis has improved. (1 = yes, 0 = no).

The Questions

1. Produce charts (or use some other method) to show the distribution of age and years since diagnosis. Briefly describe what those charts show you about the distribution of the variables.
2. The people should have been randomly allocated to each of the two groups (treatment and control). Carry out a test, to compare the mean age of the people in each group. What is the result? What do you conclude?

   Carry out a test to compare the proportions of males and females in each group. What is the result? What do you conclude?
3. Was the treatment successful at reducing the PASI?
4. Did a higher proportion of people in the treatment group feel that they had improved than those in the control group?

Notes

1. You should check for outliers and distributional assumptions. If you find severe outliers, you can assume that these were data entry errors, and act appropriately.
2. Answers should be brief. For example, do not spend time describing what you did in SPSS, just provide the answer. You should aim to produce something that appears similar in style to something that would appear in a medical journal, such as the BMJ (have a look at www.bmj.com).
3. You should use 0.05 as the significance level cut-off, and 95% confidence intervals. Where appropriate, present both p-values and CIs.
4. Think very hard before you cut and paste from SPSS output. This is almost always inappropriate. You will lose marks for inappropriate material.
5. People do badly in these assignments, because they do inappropriate statistics. Read your answers, and determine whether they really answer the question.
6. People don't do well in these assignments because they write too much. They worry that they must do the correct thing, and to ensure that they have done the correct thing, they do everything that they can think of. Although this increases the probability that they have done the correct thing, it means that they have certainly done a lot of things which were not correct.