Experimental Design 2004-5
Problem Sheet

A study has shown that students who do not attend lectures are seven times more likely to fail a particular course than those who do. Clearly teaching method has an impact on the pass rate.

There are different ways of teaching any subject. For example:

A: lectures

B: zetetic methods, in which the students simply tackle progressively harder problems

C: reading

D: computer assisted learning

Of course, we could combine these, but for now let's assume we simply want to compare methods A, B, C, and D.

Our aim is to design a study to see which is the best teaching method. Suppose we have 100 students, and we are at liberty to teach each of them by any method. After teaching them, we then test them by giving them all the same examination.

1. The management consultant carrying out the study decides that (i) it is important that exactly 25 students should be assigned to each method, and (ii) asks the students to allocate themselves to each method.

   (a) Is he right in (i)? What would be the effect of allocating different numbers to the methods?

   (b) (ii) is a fundamental error. Describe why and give an example of why it could lead to seriously mistaken conclusions.

   (c) How might you assign the students to the methods? Why is your strategy a good idea - what problems does it protect against?

2. It is obvious that all students on method A are unlikely to obtain exactly the same mark. The same applies to the other treatments. List four possible reasons why we might expect the marks to differ (for example (and I expect four others), perhaps the commitment of the student might influence things). Two of these should be influences which we can control, and two should be influences which we cannot.

3. We can see which method is best by comparing the means of the scores on each group. But the mean of a sample is a random variable. The mean for group A, for example, would be different if we repeated the experiment with a different set of students. We need to be confident that our estimate of the mean is accurate:

   (a) Why?

   (b) What is a general method for reducing the variance of a mean?

   (c) For the influences you gave in your answer to question 2 which we cannot control, how might we design the experiment so as to minimise their impact on our results?

4. In the above study 'teaching method' is one factor. Suppose that 'teacher' is another factor, with two levels T1 and T2. Suppose also that each teacher takes four classes, each using one of methods A, B, C, and D, so that we have a two factor experiment.

   (a) What would an 'interaction' mean here?

   (b) How would we detect such an interaction?

   (c) What would an interaction imply for our conclusions?

5. With the two factor experiment described in 4, we could have at most 12.5 students in each factor combination, on average. This is not very many, so that the results might be unreliable. To increase the numbers, the experiment could be repeated the year after.

   (a) What further complications might this cause?

   (b) How might you design the experiment to control such problems?

6. An alternative to the experiment discussed above would simply be to look back at students who had been taught by the different methods, in previous years. Studies where there is no purposive design, but where one simply analyses data which have already arisen, are called observational studies. Almost all data mining studies are observational. Such studies have major disadvantages compared with experimental studies. Describe some.

7. Experimental design schools such as the Deming school and Six Sigma stress methods for the reduction in variation between manufactured products. Why is this desirable?