Experimental Design 2004-5
Solutions

1. (a) No, it is not necessary to have equal numbers of students in each group. The analysis is based on comparing the means in the groups. However, greatest accuracy in this comparison, and the most powerful tests, are obtained when the groups are equal. Thus, it is not necessary to have equal numbers, but it is desirable.

   (b) It is entirely possible that different types of students will choose different methods. We will then be unable to say if any observed difference between the means is due to the different teaching methods or these other differences between the students.

   (c) Random allocation of the students to treatments would be a good idea. It makes it unlikely that a dramatically unbalanced assignment will occur (e.g. all the older students in one group). It also allows us to make proper statistical probability statements, and carry out statistical tests.

2. We can control things such as the time students spend on each method, the extent to which students are required to turn up to sessions (e.g. it would probably not be a good idea to require students to attend lectures in A, but just let them read in their own time in C. Any differences would then be confounded with the amount of time spent.). We cannot control things such as the ages of the students, their previous experience of similar material, etc. We must design our experiment to allow for such things.

3. (a) If the estimates of the means are not accurate, then the estimates of differences between the means will not be accurate.

   (b) The variance of a mean can be reduced by using a larger sample.

   (c) For things we cannot control, we should either block the students, so that the younger ones are spread evenly across the treatments, for example, or we should randomise the allocation.

4. (a) An interaction would mean that the size (and possibly the sign) of the difference between two teaching styles depended on who was teaching.

   (b) We would detect such an interaction and compare it with its standard deviation using a t-test. The null hypothesis is one of no interaction, and the test would tell us how often an interaction as large as that we observed would be observed by chance if this null hypothesis is true. If this was very small, then we would have evidence favouring the existence of a non-zero interaction.

   (c) It would make it very difficult to draw general conclusions about which teaching method was best.

5. (a) Complications arise because the circumstances are likely to change between years. The student population might have changed - perhaps tougher entry grades were required, perhaps larger fees put less able students off, etc. The teachers might be different.

   (b) One would want to try to use the same teachers, and to match on student type. However, no amount of statistical adjustment can correct problems arising from poor data, so one should not expect miracles. The best one can do is think of very possible influence one can, and design one's experiment to allow for them.

6. It is likely that the circumstances under which old data were collected will be different from those relating to today. Perhaps, for example, the teachers were different.

7. A high degree of variation is likely to mean many products are too poor, and have to be rejected, or many instance of a service are not up to scratch (dissatisfied customer). It is important that the values are tightly clustered about the mean (small variance) so that we know our conclusions about which means are larger apply generally to the products, not merely the means. For such reasons, experimental design courses in management often stress controlling the variation as well as the mean values