Chapter 4 Between-individual variation, replication, and sampling

Answers to additional self-test questions

Q4.1 Any variation in the response variable between individuals in our sample (between-individual variation) that cannot be attributed to the independent factors is called by a variety of different names: random variation, inherent variation, background variation, extraneous variation, within-treatment variation, or noise. Which is your favourite and why?

Nick likes 'within-treatment' variation because it really makes clear that we are focussing on variation that cannot be explained by the factors we are interested in. Graeme likes 'noise' because it requires least typing and is easier to spell! ☺

Q4.2 Explain why replication is an important part of experimental design.

Replication is a way of dealing with the between-individual variation due to the random variation that will be present in any life science experiment. That is, our subjects will differ not just potentially because of factors that we are interested in, but also for other reasons that we are not interested in and don't even necessarily understand. The more replicates we have, the greater the confidence we have that any difference we see between groups is due (at least in part) to the factors that we are interested in and not due to these nuisance factors.

Q4.3 In the example in the book looking at height difference between ten-year-old Scottish boys and girls, how big a difference might you expect, and what would you expect the standard deviation of the values for one particular sex to be?

If there is a difference, we would expect it to be a small number of centimetres, 15 at most. If the heights of one sex are normally distributed, then we would expect most of the spread of individuals to be in a range spanning four standard deviations. We might imagine that almost all individuals of one sex might be contained within a spread of 25cm, giving a standard deviation of around 6cm.

Q4.4 Can you give an example of cluster sampling in a study not involving humans?

If we were exploring the sexual variation in activity patterns of overwintering wildfowl in Scotland, we could randomly select water bodies to investigate, and at each water body monitor the activity of all or some of the male and female birds present.

Q4.5 Can you give an example of convenience sampling involving non-human subjects?

If we wanted to survey the prevalence of parasites in hedgehogs in Scotland then we might ask our colleagues to collect any road-killed individuals they come across and bring them to us.

Q4.6 Assume that we want to compare the mass of field mice in Greater Glasgow (a city in the west of Scotland) to those on Great Cumbrae (an island near the west coast of Scotland). How would you go about sampling?

The standard way to survey small rodents is live-trapping them using something like a Longworth trap. We are anxious that we get a representative sample of animals. To do this we should mark animals that we trap to avoid recording the same animal twice, and we should move our traps around the survey areas, making sure that we are unbiased with respect to habitat type. We would make sure we take the trouble to climb to the tops of hills and go to parts of Glasgow that are unfamiliar to us. We would also make sure that we carried out sampling simultaneously in the two areas to avoid time being a confounding factor. Some mice are more likely to enter a trap than others and this might be related to mass (if juveniles are, for example, less wary), however this bias should only affect our results if the effect of mass on likelihood of entering an encountered trap is different between the two sites. A mechanism that would cause this is not obvious, but it could be explored by filming a subset of traps at both sites.

Q4.7 In the mice study above, can you think of any biological reasons why you might expect a difference?

It could be that there are genetic differences between the populations due to a founder effect. It could also be that there are different selection pressures on mice in the two environments due to climate differences, habitat differences, differences in predation pressure, or differences in pressure from competitors. It is commonplace for mice on islands to be larger than the same species on nearby mainland.

Q4.8 We want to describe the average height of third year undergraduates at the University of St Andrews (where Graeme works). Would the third year Zoology class be a reasonable sample to use?

No; using a group like this would be OK if an individual's degree choice were independent of factors likely to influence height, such as sex, socioeconomic group, or ethic group. However, all of these factors are likely to be linked to degree choice and all of these factors are also linked to height.

Q4.9 Is there any aspect of third year undergraduates at the University of St Andrews for which the third year Zoologists might be a reasonable sample to use?

Certainly for things like 'how you spend the summer' or 'what you plan to do after university' the Zoologists would be a hopelessly biased sample. Some things, like blood group, or views of the quality of the university's sports facilities, or preference for Coke versus Pepsi, might be less problematic, but again it does not make sense to take a chance. When you can obtain a random sample (like this) then you should.

Q4.10 Would you expect greater variation in height in a sample of male undergraduates at the University of St Andrews, or in a sample of males frequenting bars in the town?

The latter, because of greater age variation amongst bar-goers. There may be a greater range of nationalities amongst students, but the age effect will probably be stronger.