What will this chapter tell me?

When we were learning to read at primary school, we used to read versions of stories by the famous storyteller Hans Christian Andersen. One of my favourites was the story of the ugly duckling. This duckling was a big ugly grey bird, so ugly that even a dog would not bite him. The poor duckling was ridiculed, ostracized and pecked by the other ducks. Eventually, it became too much for him and he flew to the swans, the royal birds, hoping that they would end his misery by killing him because he was so ugly. Still, life sometimes throws up surprises and as he stared into the water, he saw not an ugly grey bird but a beautiful swan. Data are much the same. Sometimes they're just big, grey and ugly and don't do any of the things that they're supposed to do. When we get data like these, we swear at them, curse them, peck them and hope that they'll fly away and be killed by the swans. Alternatively, we can try to force our data into becoming beautiful swans. That's what this chapter is all about: trying to make an ugly duckling of a data set turn into a swan. Be careful what you wish your data to be, though: a swan can break your arm.[1]

[1] Although it is theoretically possible, apparently you'd have to be weak boned, and swans are nice and wouldn't do that sort of thing.

When to use non-parametric tests

We discovered in the last chapter that there are many things that can bias the conclusions from a statistical model. We also looked at several ways to reduce this bias. Sometimes, however, no matter how hard you try, you will find that you can't correct the problems in your data. This is a particular problem if you have small samples and can't, therefore, rely on the central limit theorem to get you out of trouble. However, there is a small family of tests that can be used to test hypotheses that don't make many of the assumptions that we looked at in the last chapter. They are called **non-parametric tests** or 'assumption-free tests' because they make fewer assumptions than the other tests that we'll look at in this book.[2] In general, you are better off trying to use a robust test than a non-parametric test, but we'll look at non-parametric tests because (1) the range of robust tests is limited in SPSS; and (2) non-parametric tests are a nice gentle way for us to look at the idea of using a statistical test to evaluate a hypothesis.

All of the tests in this chapter overcome the problem of the shape of the distribution of scores by **ranking** the data: that is, finding the lowest score and giving it a rank of 1, then finding the next highest score and giving it a rank of 2, and so on. This process results in high scores being represented by large ranks, and low scores being represented by small ranks. The analysis is then carried out on the ranks rather than the actual data. By using the ranks we eliminate the effect of outliers: imagine you have 20 data points and the two highest scores are 30 and 60 (a difference of 30); these scores will have ranks of 19 and 20 (a difference of 1). In much the same way, ranking irons out problems with skew. Some people believe that non-parametric tests have less power than their parametric counterparts, but this is not always true (Jane Superbrain Box 6.1). In this chapter, we'll look at carrying out and interpreting four of the most common non-parametric procedures: the Mann–Whitney test, the Wilcoxon signed-rank test, Friedman's test and the Kruskal–Wallis test.

---

[2] Some people might tell you that non-parametric tests are 'distribution-free tests' because they make *no* assumptions about the distribution of the data. However, they *do* make distributional assumptions but just not normality: the ones in this chapter, for example, all assume a continuous distribution.