Designing your own correlational research

Earlier in this chapter, we defined 'correlation' as the degree of relationship between the two sets of scores. We also showed that such correlation coefficients can be either positive (if the numbers are related in the same direction) or negative (if they are related in opposite directions). In designing studies that use correlation coefficients, you must recognize two things: (a) the sign (i.e. positive or negative) and magnitude of a correlation coefficient are two different things, and (b) correlation coefficients only make sense if the assumptions underlying them are met.

Magnitude and sign of correlation coefficients

Positive correlation coefficients will sometimes make sense and at other times negative ones will. For example, even without calculating a correlation coefficient, you would expect some degree of correlation in any group of people between height and weight (i.e. the taller people are, the more they will probably weigh, and conversely, the shorter they are the less they are likely to weigh). Other similar examples might include relationships that you would expect between waist measurements and weight or between age and weight. In the realm of language teaching, you would probably expect some degree of relationship among your students between length of language study and proficiency in that language or between whether or not students have lived in the country where a language is spoken and proficiency in that language.

Negative correlations also occur in logical ways. For instance, because of various cultural factors, the number of children per family tends to be negatively related to the family income, i.e. as family incomes goes up, the number of children per family goes down and vice versa. Similarly, it would be reasonable to expect a negative relationship between worker efficiency ratings and number of hours worked, efficiency going down as workers spend more hours working and getting increasingly tired. (A personally more discouraging example was provided by some of our younger students who, observing us as older professors, suggested that there might be some negative correlation between age and ability to remember things.) In the realm of language teaching, you might expect a negative correlation between pronunciation test scores and the age at which students began studying the language (especially if some started as very young children). You might be surprised to find a negative correlation between the scores on a very detailed English grammar test and the number of years students have spent in the country where the language is spoken. Nevertheless one of the authors of this book found just such a correlation in Japan. His interpretation was that students still in Japan pay a great deal of attention to grammar and thus score well on a picky grammar test, while students who have actually acquired English in an English-speaking country haven't studied detailed grammar rules for years and so have trouble with them. Thus, in some cases, negative correlations are a natural consequence of the way some variables are counted and how they are related with other variables, while in other cases, negative correlations may present a surprise that you will have to try to interpret.

It is important to recognize that negative correlation coefficients are not lower than positive ones. The magnitude of the correlation, whether positive or negative, tells you the degree of relationship between the two sets of numbers, and the sign tells you the direction of the relationship. If you plan and understand your study properly, you should know well before you gather data which of your correlations are likely to be positive and which negative.

Assumptions underlying correlation coefficients

As Brown put it, '*Assumptions* are preconditions that are necessary for accurate application of a particular statistical test' (Brown 1992a: 639). Furthermore, as a researcher, you are responsible for planning and actually checking these assumptions before reporting the results of your statistical analyses. Thus ASSUMPTIONS are preconditions that must be checked in advance so you can accurately and responsibly apply a particular statistic.

Assumptions underlying the Pearson r

The assumptions underlying the Pearson *r* are as follows:

1 SCALES ASSUMPTION: both sets of numbers must be continuous scales.
2 INDEPENDENCE ASSUMPTION: the pairs of numbers within a data set must be independent of one another, i.e. each number in each pair must have been generated in ways which could not have influenced the generation of the other. (See below for examples.)
3 NORMALITY ASSUMPTION: both distributions must be normal, as defined in Chapter 5 (page 136).
4 LINEARITY ASSUMPTION: The two sets of numbers, if plotted on a SCATTERPLOT or graph, should be more or less in a line.

If you find all four assumptions are met, you can reasonably interpret the resulting correlation coefficient as the degree to which the two sets of numbers are associated. If you find that your data violate one or more of the assumptions, the resulting correlation coefficient may be an overestimate or underestimate of the true state of affairs in the group you are studying.

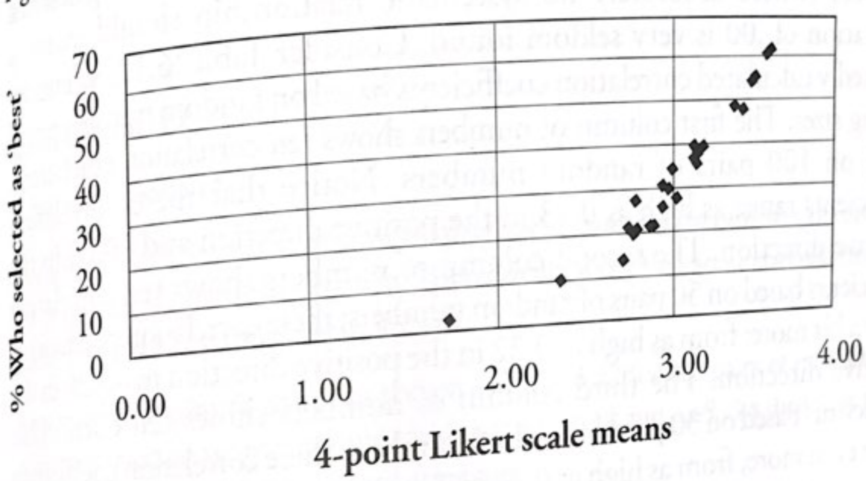
Check Assumption 1 (scales assumption) by examining each scale and confirming that each is continuous (according to the definition given earlier in this chapter).

Check Assumption 2 (independence assumption) (a) by making sure the students do not cheat or otherwise collaborate while you are testing or otherwise measuring them and (b) by examining the sets of numbers to make sure the same student does not appear twice in the data (i.e. did not generate two pairs of scores).

Check Assumption 3 (normality assumption) (a) by examining the descriptive statistics for normality, or normal distribution, (by examining them to see if you can fit at least two standard deviations on each side of the mean within the range of scores and/or (b) by creating a histogram or bar graph of the scores and deciding if it looks approximately normal.

Check Assumption 4 (linearity assumption) by plotting the two sets of numbers in a scatterplot and looking at them. For instance, if we plot the two sets of numbers that served as the basis of the demonstration calculations for the Pearson *r* (the *r* that turned out to be .93), we would get the SCATTERPLOT shown in Figure 6.1. Recall that we are examining the correlation between the Likert scale means for each item on Willing's questionnaire and the percentage who selected each item as 'best'. In the scatterplot the percentages are on the vertical axis and the Likert scale means are on the horizontal axis. Each item's pair of data points (percentage 'best' and Likert mean) in Table 6.9 is represented by a single diamond in Figure 6.1: from item 20 with 3.54 and 62, and item 11 with 3.51 and 61 all the way down to item 3 with 2.35 and 10 and item 13 with 1.69 and 3. Since the diamonds are more or less in a straight line (with the exception of that single student to the lower left, who makes the data look slightly curved), we can say that the two sets of numbers are approximately linear in relationship. If there were a marked curve in the data or if there was no clear pattern, this assumption would not be met. In such cases, correlations become very difficult to interpret.

Figure 6.1 Scatterplot to check linearity



Assumptions underlying the Spearman rho, and phi coefficients

The other correlation coefficients covered in this chapter require fewer assumptions than the Pearson *r*. Spearman ρ assumes that the scales are rank-orders (for example, parent and child movie preferences) and requires independence and linearity. The phi coefficient (Φ) assumes that both scales are nominal (i.e. are named categories, with two possibilities each, like male and female), are independent of one another (for example, are not the same people repeated), and are in a linear relationship.

Interpreting correlational research

In earlier sections of this chapter, we showed how three types of correlation coefficient can be calculated: the Pearson *r*, Spearman *ρ*, and the phi coefficient (*Φ*). In all three cases, those coefficients are useful for understanding the degree of relationship between the numbers involved. However, certain cautions must be observed in interpreting such coefficients: even random numbers can show some degree of correlation; statistical significance does not imply meaningfulness; and correlation coefficients do not indicate causality.

Even random numbers can show some degree of correlation

One problem that was noticed early in the development of correlation coefficients was that they seldom turned out to be exactly zero. In fact, even when calculating a correlation coefficient based on randomly generated numbers (where absolutely no systematic relationship should exist), a correlation of .00 is very seldom found. Consider Table 6.12, where we repeatedly calculated correlation coefficients based on random number sets of varying sizes. The first column of numbers shows ten correlation coefficients based on 100 pairs of random numbers. Notice that these correlation coefficients range as high as 0.13 in the positive direction and –0.08 in the negative direction. The second column of numbers shows ten correlation coefficients based on 50 pairs of random numbers; these correlation coefficients range a bit more: from as high as 0.22 in the positive direction to –0.26 in the negative direction. The third column of numbers shows ten correlation coefficients based on 30 pairs of random numbers; these correlation coefficients range even more, from as high as 0.34 in the positive direction to –0.38 in the negative direction. The fourth column of numbers shows ten correlation coefficients based on ten pairs of random numbers; these correlation coefficients range still more, from as high as 0.48 in the positive direction to –0.49 in the negative direction. Finally, the last column shows ten correlation coefficients based on five pairs of random numbers; these correlation coefficients range the most: from a staggering 0.92 in the positive direction to an even larger –0.93 in the negative direction.