Chapter 3

3.1

There is a split in psychological research between ‘experimental’ psychologists, who use techniques such as ANOVA to examine mean differences, and ‘correlational’ psychologists who analyse their data using regression- and correlation-based statistical techniques. Both these techniques do the same thing, but in a different way.

3.2

Regression can be used to analyse data from an experiment about a categorical independent variable, in the same manner that a t-test can. This becomes quite clear when we take a closer look at the outcomes of a regression analysis in this scenario.

The data in such an experiment consist of two groups, a research- and control group, representing the presence and absence of the independent variable, respectively. The difference in the value for the dependent variable between the two groups is of interest here (as it shows the effect of the independent variable). The beta-coefficient (also known as the slope) of the regression analysis here signifies the change in value between the two groups. Thus the slope of the regression shows the difference between the groups; more specifically the difference between the means. (Also: the intercept is the mean for the control group).

A regression analysis also gives a t-value and significance in its output. These are equal to the t-value and p-value given in a t-test of the same data.

3.3

Regression can also be used to analyse data in the same manner that a one-way ANOVA can. This analysis does become a bit more complicated, however, as data for an ANOVA has not 2 but (at least) 3 groups. Thus these groups (representing variations of the independent variable) need to be recoded into a series of variables in order to be correctly analysed in a regression analysis. Two ways of doing this are:

Dummy coding: Here one group (the control group) is set as the reference group. New dummy variables are then created to signify the condition of the other groups. Practically this means that two independent variables are created to signify membership of the groups (called group1 and group2 for example). Values of the control group are signified by a 0 in as corresponding value in the dataset of these two independent variables. Values of group 1 correspond to a 1 in the group1 variable and a 0 in the group2 variable. Accordingly values of group 2 correspond to a 0 in the group1 variable and a 0 in the group2 variable. (see table 3.9).

Effect coding: When there is no reference group, effect coding is used. As there is no reference group, there is no group that can be given a 0 value across all independent variables. So it would seem an independent variable must be created for each group (in the same manner as in dummy coding).

However, this would result in the number of variables being equal to the number of groups, and the regression analysis would not work out correctly. To combat this a random group is chosen and signified with a -1 value across the independent variables (this group then doesn't have its own independent variable). (see table 3.13).

After such coding a regression analysis will result in the same table typically produced by an ANOVA test. After an ANOVA the nature of differences is usually understood by interpreting so-called 'post hoc' tests. Something similar can be done by examining the beta-coefficients (slopes) of each independent variable.

All this shows that regression can do anything that a one-way ANOVA can. It is important to know this because this version of regression analysis works better than the ANOVA in certain circumstances, particularly when it comes to analysing change scores.

It must be taken into account, however, that such testing is not corrected for multiple comparisons, and thus the chance of a type 1 error, the chance of falsely rejecting the null hypothesis, increases. Luckily there are corrections that can be applied to fix this discrepancy, the simplest being the Bonferroni correction. This correction is done by dividing the original significance level by the number of comparisons made.