

Try to work out the probability of rolling an even number (the answer is in the ‘Answers’ section of the book).

4.1.1 Conditional probabilities

For the purpose of research in psychology, we need an understanding not only of probability but also of *conditional probability*. A conditional probability is the probability of some event taking place, which is dependent upon something else. For example, the probability of Arsenal winning a cup final this year might be 70% if they had all their players fit, but might be only 30% if they had most of their strikers injured. These are conditional probabilities in that they are dependent upon the fitness of the Arsenal team. Another example of a conditional probability is the probability of someone buying this statistics book. Ordinarily, given that there are probably at least 99 other texts on the market, the probability of someone buying this book would be about 1 in 100 (or 1%). If a tutor recommends it, however, then the probability may change to about 1 in 5 (or 20%). This latter probability is a conditional probability; it is the probability of someone buying the book conditional upon a tutor recommending it. We mentioned two examples of conditional probabilities earlier in the chapter. The probability of contracting cancer if you smoke cigarettes is a conditional probability, as is the probability of coronary heart disease if you drink moderate amounts of beer. Try to ensure that you understand what conditional probabilities are, as you will come across them more in Chapter 5 when we explain hypothesis testing.

Definition

A *conditional probability* is the probability of a particular event happening if another event (or set of conditions) has also happened.

Activity 4.3

Which of the following are conditional probabilities?

- (a) The probability of being struck by lightning while playing golf
- (b) The probability of winning the Lottery
- (c) The probability of winning an Olympic gold medal if you do no training
- (d) The probability of getting lung cancer if you smoke
- (e) The probability of manned flight to Mars within the next ten years
- (f) The probability of having coronary heart disease if you drink moderate levels of beer

4.1.2 Applying probabilities to data analyses: inferential statistics

Inferential statistics are techniques employed to draw conclusions from your data. When we conduct research we typically want to draw some conclusions about what we are observing: that is, we want to make inferences. The reason why we investigate the relationship between mental health and procrastination or that between cigarette smoking and short-term memory is to understand them better. Similarly, the reason why we would conduct a study on people who eat chocolate sponge with tomato ketchup at 6.30 in the morning is that we want to know why on earth they do it. To answer such questions we need to draw conclusions from our data.

Given the following description of events, what conclusion are you likely to draw?

On a busy afternoon in the centre of Wolverhampton, a man was seen sprinting around a corner and running along the High Street. He was obviously in a hurry and did not care that he knocked people out of his way. About three seconds later a policewoman also came running around the corner and up the High Street.

One likely conclusion from this description is that the policewoman is trying to catch up with and arrest the person running in front of her. In the majority of cases this would be a reasonable conclusion to draw from the description above. However, it could be that the policewoman is following a plain clothes officer and they are both rushing to the scene of a crime.

You can see how easy it is to jump to the wrong conclusion in real-life events. The same sorts of error can arise in psychological research because we are trying to draw conclusions from statistical analyses. Remember, when we test a sample of individuals we are generally doing so to enable us to draw conclusions about the population from which the sample was drawn. If we wanted to find out whether mental health was related to procrastination, we would randomly select a sample of individuals and get measures of these two variables. From these data we would then try to make some inferences about the relationship between the two variables in the population. This is what we use inferential statistical techniques for. It is, however, possible that we may draw the wrong conclusions from our statistical analyses. This is because the statistical techniques we use in order to draw conclusions about underlying populations are based upon probabilities. We therefore need to be constantly aware of the fallibility of such techniques.

Example from the literature

Mental health procrastination

Stead, Shanaham and Neufeld (2010) used inferential statistical techniques to assess the relationship between mental health and procrastination. The researchers asked 200 students to complete a number of online questionnaires relating to mental health, stress and procrastination. Using a complex correlational analytical technique for their inferential statistics, they found that procrastination and stress were both related to poorer mental health. These statistical techniques enabled Stead *et al.* to argue that they had support for their hypothesis that procrastination is related to poorer mental health.

4.2 The standard normal distribution

We have explained what we mean by probability and also, in Chapter 3, described to you the family of distributions known as the normal distributions. We would now like to explain an even more important distribution known as the *standard normal distribution* (see Figure 4.1). As is indicated in Figure 4.1, the standard normal distribution (SND) is a normally shaped distribution with a mean of zero and a standard deviation of 1. Because of these and other characteristics, the SND is a very useful distribution. Using this distribution we are able to compare scores from different samples, compare different scores from the same samples and much, much more.

In order to use the standard normal distribution for analysing our data, we need to transform the scores in our samples to the standard normal scores. This is achieved by subtracting the mean from each score and then dividing by the standard deviation. The result is called a *z-score*. The z-score is expressed in standard deviation units: that is, the z-score tells us