1 Introduction

What this book tries to do

My aim is to help you to design, carry out, analyse and interpret simple experiments. The focus is on the use of the experimental approach when studying people and their ways, particularly as carried out by psychologists – although the same principles apply to experiments carried out by other social or behavioural scientists, and professionals in applied fields such as education, social services or health care. As far as statistics and statistical tests are concerned, the book is a 'cook-book'. It has been designed to guide you to the right recipe – the statistical test appropriate to the problem – and then carry you through the steps of the recipe with a maximum of detail. The intention is not to train statisticians, but to give people doing simple experiments a range of useful statistical tools. Formulae are not derived, and the only mathematical requirement is some elementary algebra, together with the ability to substitute numbers into formulae. As mentioned in the preface to this edition, it is now possible to use software available on most microcomputers to actually carry out the computation. If you have a large amount of data only a masochist would do the analysis by hand. However, it does help to know what you are doing and there is much to be said for carrying out the analysis of simple experiments with relatively small amounts of data 'by hand' (in practice this means using a simple electronic calculator to do the multiplying, squaring, etc.).

However, the book is not just a 'cook-book'. Apart from the fact that such a compendium of statistical recipes would be deadly boring and hardly likely to kindle any flames of interest or

enthusiasm, it would not be able to fulfil the stated aim of the book. If you are to select a particular statistical test, you need to appreciate what the test is capable of doing, why it is that that particular test is appropriate and others are not. In order to achieve this, some kind of understanding, even if only at an intuitive level, is essential. Hence an attempt is made to talk around and lead up to the tests in such a way that that kind of understanding has a chance to develop.

There are many sad stories of students, burning to carry out an experimental project, who end up with a completely unanalysable mishmash of data. They wanted to get on with it and thought that they could leave thoughts of analysis until after the experiment. They were wrong. Statistical analysis and experimental design must be considered together and, whilst there are broad principles of experimental design which will be covered, they cannot easily be reduced to recipes guaranteed for every eventuality.

Using statistics is no insurance against producing rubbish. Badly used, misapplied statistics simply allow one to produce quantitative rubbish rather than qualitative rubbish.

Limitations of the book

Do not get the impression that all the techniques of statistical analysis are to be laid at your feet in the next chapters. This is not the case. You will be given a severely limited range of techniques. What I do hope, however, is that you will be able to take research questions that you are interested in and turn them into simple experiments which can be analysed meaningfully by one or more tests out of this repertoire.

This will not always be possible and you may well find that in trying to fit your experiment into the strait-jacket of these techniques, so much violence is done to the original idea that it is not worth doing.

The intention is that you should be able to recognize situations like this. Your strategy then is to try to approach the research question from a different direction. It may be that, for one or more of several possible reasons, your question is just not amenable to dealing with by an experiment. We will see later in the chapter that

while there are very good reasons for seeking to use experimental approaches, it does require rather special circumstances for this to be a sensible strategy. And even if an experiment were sensible, it may well be that your questions call for more complex designs or statistical analyses than those covered here. However, do not be discouraged; it is surprising how many interesting questions can be investigated in simply designed experiments for which the simple designs and straightforward statistical tests described here are adequate and appropriate.

Requirements from the reader

The main requirements, virtually the only requirement, is that *you should want to do experiments.* If you are certain that you don't want to do experiments then you are wasting your time reading this book and should find a more profitable way of spending your time. It is true that a knowledge of statistics and experimental design enables one to understand and evaluate other people's experiments, and this should come as a bonus from reading the book. But even if that is your aim, the experience of having carried out your own experiments will stand you in good stead.

This book is addressed directly to the reader who wants to experiment but doesn't know how to go about it (including those who have started out with that intention but have got confused or dispirited). My interest and sympathies lie with you and your concerns. If you already have a specific problem or question that you would like to turn into an experiment, well and good. It will not be a bad thing for you to go through, considering each of the approaches discussed in turn to see whether it fits in with your research question. If you do not have a problem, no need to worry. Providing you keep your eyes and ears open and start thinking 'experimentally' you will be assailed at each and every turn by problems that might form the basis of interesting experiments. If, by the time you have reached the end of the book, you are still keen but remain without such a problem, the last chapter suggests some of the things that you might do about it.

What is an experiment anyway?

In everyday use the terms 'experiment' and 'experimental' are used very generally to refer to some kind of trial or investigation. Here, the meaning is much more specific.

In the simplest experiment, one investigates the relationship between two things by deliberately producing a change in one of them and looking at, observing, the change in the other. These 'things' in which change takes place are called **variables**.

The variable which we, as experimenters, are directly manipulating is called the **independent variable**. The variable in which we are looking for any consequent changes is called the **dependent variable**. The independent variable is so-called because it is independent of what the participants in the experiment actually do (it is predetermined by the experimenter). The dependent variable is so-called because changes in it are (potentially, at least) dependent on changes in the independent variable. To take an example: suppose that one is interested in the effect of financial reward on the performance of some complex task. An experiment could be devised where the independent variable would be the amount of money given, and the dependent variable some measure of the performance of the task, for example, the number of errors made.

A second important feature of the experiment is that the experimenter not only deliberately manipulates the independent variable and looks for possible changes on the dependent variable, but also seeks to control other variables so that they do not affect the outcome. Suppose that the problem or research question is concerned with the relative effectiveness of lectures given in the early morning or late afternoon. In this case the independent variable is the time of the lectures, and the dependent variable some measure of the effectiveness of the lectures. This measure might, for instance, be the amount of information remembered about the lectures. There are many other things which might affect the amount remembered, apart from the time of the lecture. The students or the lecturer might be different (different actual people or different in terms of their performance – are you as alert in the afternoons as in the mornings?). Or the material of the lecture. Or the lecture room. Or the conditions in the room. Or what happened in the

hour before the lecture. One could go on, and on, and on . . . What we are saying here is that there are other variables apart from the independent variable, which might affect the dependent variable. Providing experimenters have sufficient patience, cunning and ingenuity, they can arrange to control all these other variables in such a way that they will not affect the assessment of the relationship between the independent and dependent variable.

The experiment is not limited to a consideration of a single, independent variable and its possible effects on a single dependent variable. Most published experiments involve the manipulation of several independent variables, and there is a trend towards studies with more than one dependent variable. However, such studies, to be adequately analysed, tend to involve complex statistical techniques which are beyond the range of the introductory text.

The big advantage of experiments

In a well-designed experiment where adequate control measures have been taken against other variables, there is an inherent plausibility to the claim that changes observed in the dependent variable have been *caused* by your manipulation of the independent variable. This ability of the experiment to get at causal relationships is its big advantage over other approaches.

With non-experimental approaches, such as the typical survey, claims about causality are more difficult to substantiate. In a situation where many variables are free to vary, you can never be sure that changes in one particular variable occur as a result of changes in a second particular variable. It is always possible that a third variable is causally related to both of the first two and has produced the relationship observed between them. For example, an observational study of the 'Trooping of the Colour' ceremony over a number of years could well demonstrate a strong relationship between the softness of the tar on the parade ground and the number of guardsmen fainting on parade. Any direct causal relationship, such as the poor chaps being overcome by noxious fumes from the tar, appears to be highly unlikely and I would surmise (although I have not tested this) that an actual experiment would fail to demonstrate any relationship. The obvious causal link is, of

course, that of temperature. The rising temperature causes both the tar to soften and the guardsmen to faint.

A more serious example is the relationship found by several surveys between smoking and lung cancer. For ethical reasons no actual experiments have been carried out with humans to investigate this link (see p. 94 for further discussion). Hence it is possible that, although there is a strong relationship between smoking and cancer, smoking may not be the cause of cancer. The cause may be, for example, in some personality or other psychological characteristics which predispose to both smoking and cancer.

Nevertheless, non-experimental studies do have an important part to play in investigations, psychological or otherwise. There are many situations, other than the smoking and lung cancer one, where direct experimentation is not possible. Perfectly respectable sciences such as astronomy and geology have to rely almost exclusively on direct observation without the possiblity of the active manipulation of an independent variable required for an experiment. Working with humans, and indeed with other animals, there are ethical reasons why variables involving painful stimuli, surgical operations and so on, should not be manipulated, and here non-experimental studies may well be preferable.

There are also strong arguments for conducting at least the initial stages of investigation into an area in an open-ended exploratory fashion, even when experimentation is feasible. In this way it is possible to get some idea of which variables are most important. The experiment is a very precise tool. As experimenter you are, as it were, putting your bets on a particular independent variable and a particular dependent variable as the ones likely to be causally related. Even in more complex experiments you are restricted to just a few of the many possible variables that you might choose. For the choice of an experiment to be a sensible strategy you need to have a justified confidence in your choice of variables, obtained either from previous work carried out by others or by exploratory work of your own.

Some difficulties with experiments

The preceding paragraphs make it clear that there are some situations and research problems where it may not be feasible nor

appropriate to carry out an experiment. For some, the time may not yet be ripe in that insufficient is known to devise a sensible design. For others, such as the lung cancer example, experiments are effectively ruled out in principle.

There are other difficulties. A common criticism of experiments is that they are artificial and over-simplified. To fulfil the requirements of control of variables you may be in danger of throwing away the baby with the bath-water. Suppose you start out with a research question on some aspect of car-driving behaviour. The process of designing an experiment, and in particular the need to control the many variables encountered in 'real' driving on the open road, may well lead you to end up with a study of persons pushing buttons in a laboratory cubicle in reponse to a flashing light.

The mere fact that an experiment is typically a rather special event taking place in a special place may mean that what is found out is of limited generality. Human participants bring their expectations and perceptions into this situation; they may be seeking to please the experimenter or the reverse. The effects of the experimenters' own expectancies on the behaviour of participants have been extensively researched. Barber (1976) provides a balanced account of these, and other, potential pitfalls.

Experiments and laboratories

Traditionally, experiments take place in laboratories. The prime reason for this is that it helps in controlling variables. The use of specialized equipment, such as the tachistoscope, is essentially so that the experimenter can obtain very precise control over what happens to participants. However, there is no iron law that says that experiments must take place in laboratories, and for certain types of question so-called 'field' experiments may well be preferable. Remember that the central feature of the experiment is the deliberate manipulation of some variable. The ingenious experimenter may well find ways of setting this up in a school, hospital ward, pub or wherever. Some lack of control over variables is probably inevitable but the increased naturalness of the setting may well reap considerable benefits.

A warning

Finally, a warning against over-optimism, and against the rejection or devaluing of findings and evidence from non-experimental approaches. There was a wave of enthusiasm for experimentation in education in the 1920s. This was followed by a wave of pessimism and disillusionment. The advocates of experimentation assumed that progress in teaching methods had been slow just because there had been little or no experimentation. When their experiments proved to be tedious, equivocal, difficult to replicate (i.e. for their findings to be supported when the study was repeated) and to accord with common sense, then disillusionment and rejection of experimentation took place.

The justification for the experimental method is not as a panacea to be used in all situations to seek answers for all problems. It is that when questions amenable to experimental 'attack' can be devised, it is the simplest and most straightforward method we know for getting at cause and effect relationships.