Introduction to

# Statistics in Psychology

with SPSS

Sixth edition

Dennis Howitt
& Duncan Cramer

ALWAYS LEARNING

PEARSON

# Introduction to Statistics in Psychology

# Introduction to Statistics in Psychology

## Sixth Edition

**Dennis Howitt** Loughborough University

**Duncan Cramer** Loughborough University

PEARSON

Harlow, England • London • New York • Boston • San Francisco • Toronto • Sydney • Auckland • Singapore • Hong Kong
Tokyo • Seoul • Taipei • New Delhi • Cape Town • São Paulo • Mexico City • Madrid • Amsterdam • Munich • Paris • Milan

**Pearson Education Limited**
Edinburgh Gate
Harlow CM20 2JE
United Kingdom
Tel: +44 (0)1279 623623
Web: www.pearson.com/uk

First published 1997 (print)
Second edition published 2000 (print)
Revised second edition 2003 (print)
Third edition 2005 (print)
Fourth edition 2008 (print)
Fifth edition 2011 (print)
**Sixth edition published 2014** (print and electronic)



ISBN: 978-1-292-00074-9 (print)
978-1-292-00076-3 (PDF)
978-1-292-00075-6 (eText)

**British Library Cataloguing-in-Publication Data**
A catalogue record for the print edition is available from the British Library

**Library of Congress Cataloging-in-Publication Data**
Howitt, Dennis.
Introduction to statistics in psychology / Dennis Howitt, Loughborough University, Duncan Cramer, Loughborough University. -- 6th Edition.
pages cm
ISBN 978-1-292-00074-9
1. Psychometrics. I. Cramer, Duncan, 1948- II. Title.
BF39.H74 2013
150.1'5195--dc23 2013037101

10 9 8 7 6 5 4 3 2 1
18 17 16 15 14

Cover image © Getty Images

Print edition typeset in 9.5/12pt Sabon by 35
Print edition printed in Great Britain by Butler, Tanner and Dennis Ltd

NOTE THAT ANY PAGE CROSS REFERENCES REFER TO THE PRINT EDITION

# Brief contents

# Contents

## Companion Website

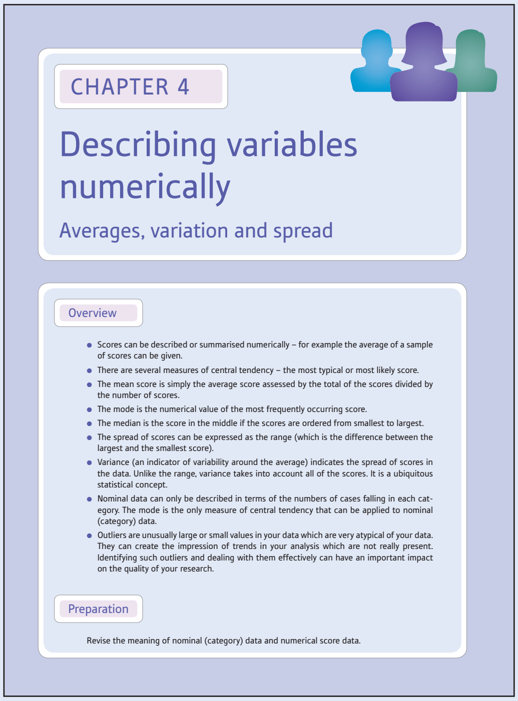
For open-access **student resources** specifically written to complement this textbook and support your learning, please visit **www.pearsoned.co.uk/howitt**

ON THE WEBSITE

## Lecturer Resources

For password-protected online resources tailored to support the use of this textbook in teaching, please visit **www.pearsoned.co.uk/howitt**

# Guided tour



CHAPTER 4

Describing variables numerically

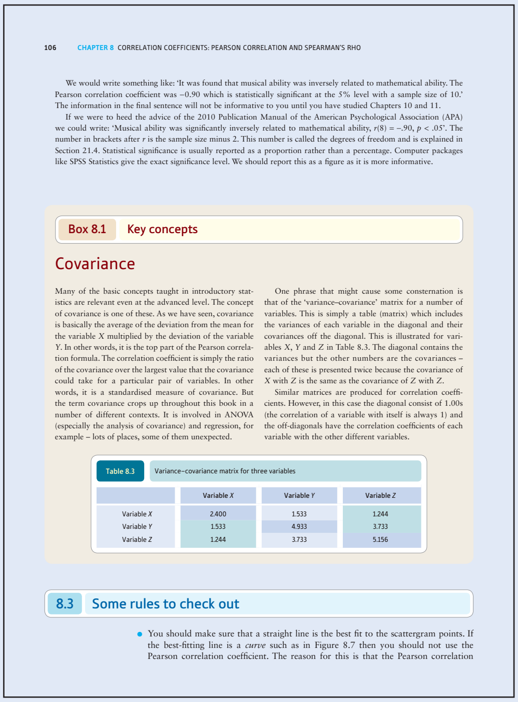
Averages, variation and spread

Overview

- Scores can be described or summarised numerically – for example the average of a sample of scores can be given.
- There are several measures of central tendency – the most typical or most likely score.
- The mean score is simply the average score assessed by the total of the scores divided by the number of scores.
- The mode is the numerical value of the most frequently occurring score.
- The median is the score in the middle if the scores are ordered from smallest to largest.
- The spread of scores can be expressed as the range (which is the difference between the largest and the smallest score).
- Variance (an indicator of variability around the average) indicates the spread of scores in the data. Unlike the range, variance takes into account all of the scores. It is a ubiquitous statistical concept.
- Nominal data can only be described in terms of the numbers of cases falling in each category. The mode is the only measure of central tendency that can be applied to nominal (category) data.
- Outliers are unusually large or small values in your data which are very atypical of your data. They can create the impression of trends in your analysis which are not really present. Identifying such outliers and dealing with them effectively can have an important impact on the quality of your research.

Preparation

Revise the meaning of nominal (category) data and numerical score data.



106 CHAPTER 8 CORRELATION COEFFICIENTS: PEARSON CORRELATION AND SPEARMAN'S RHO

We would write something like: 'It was found that musical ability was inversely related to mathematical ability. The Pearson correlation coefficient was −0.90 which is statistically significant at the 5% level with a sample size of 10.' The information in the final sentence will not be informative to you until you have studied Chapters 10 and 11.

If we were to heed the advice of the 2010 Publication Manual of the American Psychological Association (APA) we could write: 'Musical ability was significantly inversely related to mathematical ability, $r(8) = -.90$, $p < .05$'. The number in brackets after $r$ is the sample size minus 2. This number is called the degrees of freedom and is explained in Section 21.4. Statistical significance is usually reported as a proportion rather than a percentage. Computer packages like SPSS Statistics give the exact significance level. We should report this as a figure as it is more informative.

Box 8.1 Key concepts

Covariance

Many of the basic concepts taught in introductory statistics are relevant even at the advanced level. The concept of covariance is one of these. As we have seen, covariance is basically the average of the deviation from the mean for the variable $X$ multiplied by the deviation of the variable $Y$. In other words, it is the top part of the Pearson correlation formula. The correlation coefficient is simply the ratio of the covariance over the largest value that the covariance could take for a particular pair of variables. In other words, it is a standardised measure of covariance. But the term covariance crops up throughout this book in a number of different contexts. It is involved in ANOVA (especially the analysis of covariance) and regression, for example – lots of places, some of them unexpected.

One phrase that might cause some consternation is that of the 'variance–covariance' matrix for a number of variables. This is simply a table (matrix) which includes the variances of each variable in the diagonal and their covariances off the diagonal. This is illustrated for variables $X$, $Y$ and $Z$ in Table 8.3. The diagonal contains the variances but the other numbers are the covariances – each of these is presented twice because the covariance of $X$ with $Z$ is the same as the covariance of $Z$ with $Z$.

Similar matrices are produced for correlation coefficients. However, in this case the diagonal consist of 1.00s (the correlation of a variable with itself is always 1) and the off-diagonals have the correlation coefficients of each variable with the other different variables.

Table 8.3 Variance–covariance matrix for three variables

| | Variable $X$ | Variable $Y$ | Variable $Z$ |
|---|---|---|---|
| Variable $X$ | 2.400 | 1.533 | 1.244 |
| Variable $Y$ | 1.533 | 4.933 | 3.733 |
| Variable $Z$ | 1.244 | 3.733 | 5.156 |

8.3 Some rules to check out

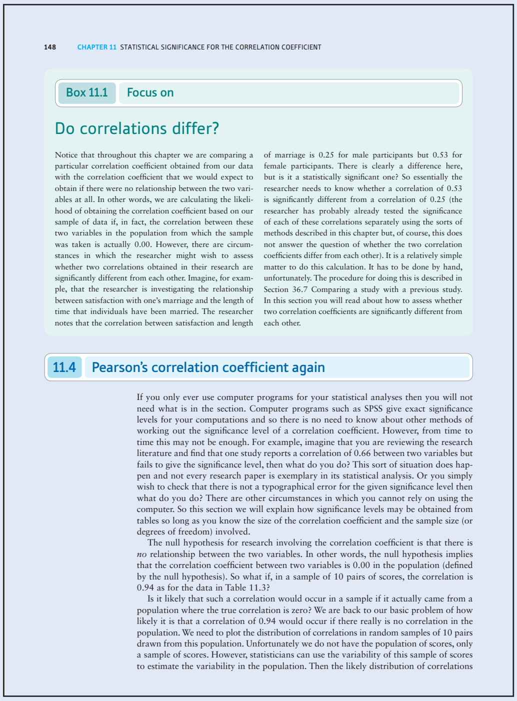
- You should make sure that a straight line is the best fit to the scattergram points. If the best-fitting line is a *curve* such as in Figure 8.7 then you should not use the Pearson correlation coefficient. The reason for this is that the Pearson correlation

## Clear overview

Introduce the chapter to give students a feel for the topics covered.

## Key concepts

Offer guidance on the important concepts and issues discussed in the text.





Box 11.1 Focus on

### Do correlations differ?

Notice that throughout this chapter we are comparing a particular correlation coefficient obtained from our data with the correlation coefficient that we would expect to obtain if there were no relationship between the two variables at all. In other words, we are calculating the likelihood of obtaining the correlation coefficient based on our sample of data if, in fact, the correlation between these two variables in the population from which the sample was taken is actually 0.00. However, there are circumstances in which the researcher might wish to assess whether two correlations obtained in their research are significantly different from each other. Imagine, for example, that the researcher is investigating the relationship between satisfaction with one's marriage and the length of time that individuals have been married. The researcher notes that the correlation between satisfaction and length of marriage is 0.25 for male participants but 0.53 for female participants. There is clearly a difference here, but is it a statistically significant one? So essentially the researcher needs to know whether a correlation of 0.53 is significantly different from a correlation of 0.25 (the researcher has probably already tested the significance of each of these correlations separately using the sorts of methods described in this chapter but, of course, this does not answer the question of whether the two correlation coefficients differ from each other). It is a relatively simple matter to do this calculation. It has to be done by hand, unfortunately. The procedure for doing this is described in Section 36.7 Comparing a study with a previous study. In this section you will read about how to assess whether two correlation coefficients are significantly different from each other.

### 11.4 Pearson's correlation coefficient again

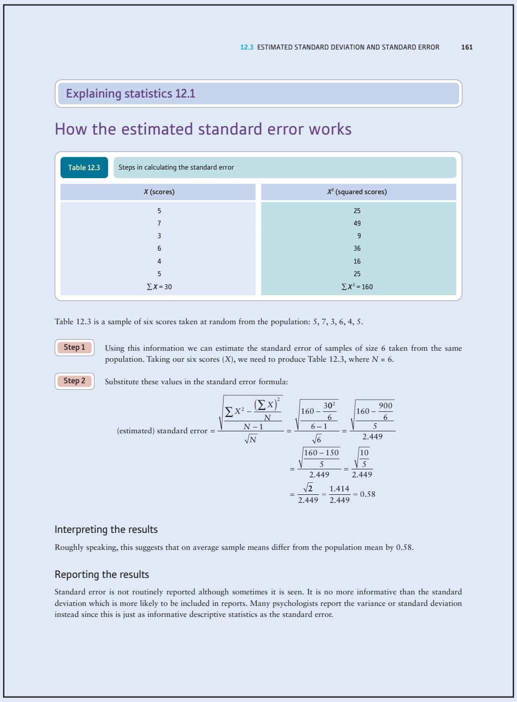
If you only ever use computer programs for your statistical analyses then you will not need what is in the section. Computer programs such as SPSS give exact significance levels for your computations and so there is no need to know about other methods of working out the significance level of a correlation coefficient. However, from time to time this may not be enough. For example, imagine that you are reviewing the research literature and find that one study reports a correlation of 0.66 between two variables but fails to give the significance level, then what do you do? This sort of situation does happen and not every research paper is exemplary in its statistical analysis. Or you simply wish to check that there is not a typographical error for the given significance level then what do you do? There are other circumstances in which you cannot rely on using the computer. So this section we will explain how significance levels may be obtained from tables so long as you know the size of the correlation coefficient and the sample size (or degrees of freedom) involved.

The null hypothesis for research involving the correlation coefficient is that there is *no* relationship between the two variables. In other words, the null hypothesis implies that the correlation coefficient between two variables is 0.00 in the population (defined by the null hypothesis). So what if, in a sample of 10 pairs of scores, the correlation is 0.94 as for the data in Table 11.3?

Is it likely that such a correlation would occur in a sample if it actually came from a population where the true correlation is zero? We are back to our basic problem of how likely it is that a correlation of 0.94 would occur if there really is no correlation in the population. We need to plot the distribution of correlations in random samples of 10 pairs drawn from this population. Unfortunately we do not have the population of scores, only a sample of scores. However, statisticians can use the variability of this sample of scores to estimate the variability in the population. Then the likely distribution of correlations

## Focus on

Explore particular concepts in more detail.





Explaining statistics 12.1

### How the estimated standard error works

Table 12.3 Steps in calculating the standard error

| $X$ (scores) | $X^2$ (squared scores) |
|---|---|
| 5 | 25 |
| 7 | 49 |
| 3 | 9 |
| 6 | 36 |
| 4 | 16 |
| 5 | 25 |
| $\sum X = 30$ | $\sum X^2 = 160$ |

Table 12.3 is a sample of six scores taken at random from the population: 5, 7, 3, 6, 4, 5.

**Step 1** Using this information we can estimate the standard error of samples of size 6 taken from the same population. Taking our six scores ($X$), we need to produce Table 12.3, where $N = 6$.

**Step 2** Substitute these values in the standard error formula:

$$\text{(estimated) standard error} = \frac{\sqrt{\dfrac{\sum X^2 - \dfrac{\left(\sum X\right)^2}{N}}{N-1}}}{\sqrt{N}} = \frac{\sqrt{\dfrac{160 - \dfrac{30^2}{6}}{6-1}}}{\sqrt{6}} = \frac{\sqrt{\dfrac{160 - \dfrac{900}{6}}{5}}}{2.449}$$

$$= \frac{\sqrt{\dfrac{160-150}{5}}}{2.449} = \frac{\sqrt{\dfrac{10}{5}}}{2.449}$$

$$= \frac{\sqrt{2}}{2.449} = \frac{1.414}{2.449} = 0.58$$

#### Interpreting the results

Roughly speaking, this suggests that on average sample means differ from the population mean by 0.58.

#### Reporting the results

Standard error is not routinely reported although sometimes it is seen. It is no more informative than the standard deviation which is more likely to be included in reports. Many psychologists report the variance or standard deviation instead since this is just as informative descriptive statistics as the standard error.

## Explaining statistics

Take students through a statistical test with a detailed step-by-step explanation.

366 CHAPTER 26 ANALYSIS OF COVARIANCE (ANCOVA): CONTROLLING FOR ADDITIONAL VARIABLES

Research examples

ANCOVA

Cumming and co-workers (2012) studied the effect of physically maturing early in adolescence on the physical activity of girls. Research has suggested that girls reduce their amounts of physical activity during adolescence and the health-related issues that this entails are obvious. Is there a role for early maturation in this? The study compared early and late maturing adolescent girls with an average age of 12.7 years. The dependent variables were health-related matters such as physical activity behaviour, physical self-concept, and health-related quality of life. In each case it was expected that early maturing girls would score lower. The analysis employed several ANCOVA analyses compared early and late maturing girls on these variables. Chronological age was included as the covariate since obviously maturation and age correlate together. Although the size of the differences tended to be small to moderate, the ANCOVAs repeatedly showed that early maturing girls scored lower on the health-related variables. It is noteworthy that early maturing girls rated themselves lower in terms of body attractiveness. This may have a bearing on their lower levels of involvement in physical activity.

Estevis, Basso and Combs (2012) investigated the effect of practice on the Wechsler Adult Intelligence Scale–IV. The participants were given the test at the start of the study and again a few months later. For some it was three months later and for the others it was six months later. They used various subscales from the test including Verbal Comprehension, Working Memory, Perceptual Reasoning and Processing Speed as well as the Full Scale IQ. They analysed the data using an ANCOVA design in which test versus retest and the various subscales were the related factors and three months versus six months was the independent factor. Gender was entered as the covariate. Bonferroni adjustment was employed to deal with the repeated significance testing problem. The interval between testing and retesting did not have a significant effect.

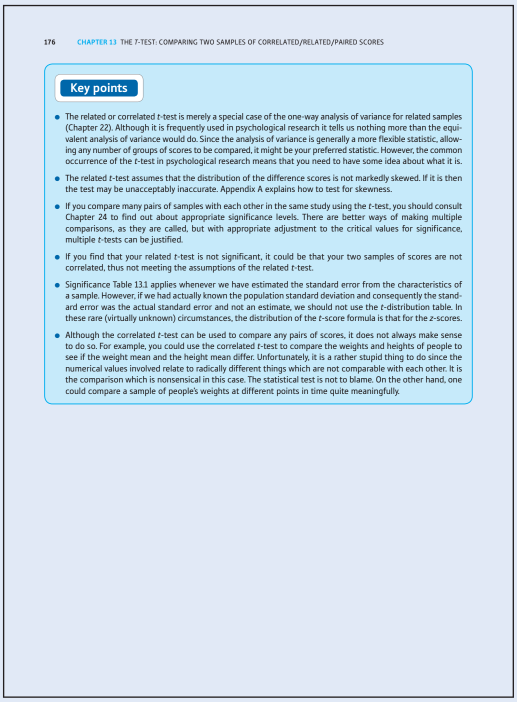
Wright and Hardie (2012) write that the previous research on the relationship between handedness and anxiety fails to indicate a clear conclusion. One reason for expecting a relationship between anxiety and handedness is that the right-hand side hemisphere of the brain is involved in negative emotional states and inhibition. Anxiety is often classified as being situational in nature or alternatively as a personality trait of the individual. The researchers found that left-handed people have statistically significantly higher scores on state anxiety which supports the idea of the role of the right hemisphere. No trait anxiety differences were found but trait and state anxiety were significantly correlated. So ANCOVA was employed with trait anxiety as the control variable because of this correlation. The handedness relationship to state anxiety remained even in this analysis. The authors suggest that left-handers are more reactive personalities and so respond with state anxiety to the new situation that they were experiencing in the research laboratory as part of the research.

Key points

- Relying on ANCOVA to deal with the problems due to employing non-randomised allocation to the cells of the ANOVA ignores the basic reason for doing randomised experiments in the first place – that the researcher does not know what unknown factors influence the outcome of the research. Random allocation to conditions is the only practical and sound way of fully controlling for variables not included in the design.
- It is not wise to use ANCOVA to try to correct for the sloppiness of your original design or procedures. Although, especially when using computers, you can include many covariates, it is best to be careful when planning your research to reduce the need for this. In randomised experiments, probably the control of the pre-test measure is the only circumstance requiring ANCOVA. Of course, there are circumstances in which pre-tests are undesirable, especially as they risk sensitising participants as to the purpose of the study or otherwise influencing the post-test measures.

## Research examples

Demonstrate how the statistical tests have been used in real research.



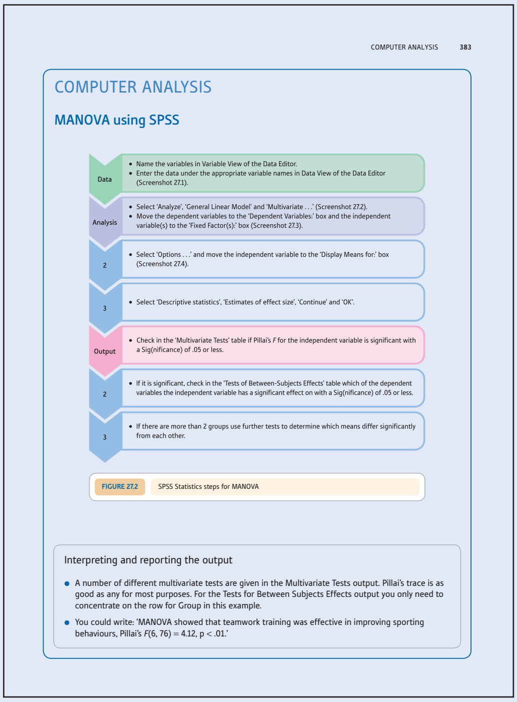
176 CHAPTER 13 THE *T*-TEST: COMPARING TWO SAMPLES OF CORRELATED/RELATED/PAIRED SCORES

Key points

- The related or correlated *t*-test is merely a special case of the one-way analysis of variance for related samples (Chapter 22). Although it is frequently used in psychological research it tells us nothing more than the equivalent analysis of variance would do. Since the analysis of variance is generally a more flexible statistic, allowing any number of groups of scores to be compared, it might be your preferred statistic. However, the common occurrence of the *t*-test in psychological research means that you need to have some idea about what it is.
- The related *t*-test assumes that the distribution of the difference scores is not markedly skewed. If it is then the test may be unacceptably inaccurate. Appendix A explains how to test for skewness.
- If you compare many pairs of samples with each other in the same study using the *t*-test, you should consult Chapter 24 to find out about appropriate significance levels. There are better ways of making multiple comparisons, as they are called, but with appropriate adjustment to the critical values for significance, multiple *t*-tests can be justified.
- If you find that your related *t*-test is not significant, it could be that your two samples of scores are not correlated, thus not meeting the assumptions of the related *t*-test.
- Significance Table 13.1 applies whenever we have estimated the standard error from the characteristics of a sample. However, if we had actually known the population standard deviation and consequently the standard error was the actual standard error and not an estimate, we should not use the *t*-distribution table. In these rare (virtually unknown) circumstances, the distribution of the *t*-score formula is that for the *z*-scores.
- Although the correlated *t*-test can be used to compare any pairs of scores, it does not always make sense to do so. For example, you could use the correlated *t*-test to compare the weights and heights of people to see if the weight mean and the height mean differ. Unfortunately, it is a rather stupid thing to do since the numerical values involved relate to radically different things which are not comparable with each other. It is the comparison which is nonsensical in this case. The statistical test is not to blame. On the other hand, one could compare a sample of people's weights at different points in time quite meaningfully.

## Key points

Each chapter concludes with a set of the key points to provide a useful reminder when revising a topic


COMPUTER ANALYSIS 383

COMPUTER ANALYSIS

MANOVA using SPSS



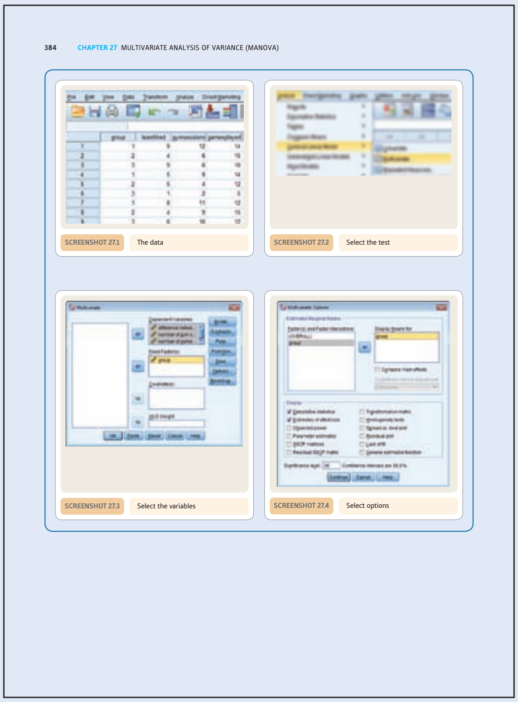
FIGURE 27.2 SPSS Statistics steps for MANOVA

Interpreting and reporting the output

- A number of different multivariate tests are given in the Multivariate Tests output. Pillai's trace is as good as any for most purposes. For the Tests for Between Subjects Effects output you only need to concentrate on the row for Group in this example.
- You could write: 'MANOVA showed that teamwork training was effective in improving sporting behaviours, Pillai's $F(6, 76) = 4.12$, $p < .01$.'

## Computer analysis

Step-by-step advice and instruction on analysing data using SPSS Statistics is provided at the end of each chapter.


384 CHAPTER 27 MULTIVARIATE ANALYSIS OF VARIANCE (MANOVA)

SCREENSHOT 27.1 The data

SCREENSHOT 27.2 Select the test

SCREENSHOT 27.3 Select the variables

SCREENSHOT 27.4 Select options

## SPSS screenshots

The guidance on how to use SPSS for each statistical test is accompanied by screenshots, so the processes can be easily followed.

# Introduction

Our hope is that this sixth edition of *Introduction to Statistics in Psychology* will contribute even more to the student learning experience. A number of changes have been made to this end. In particular, a new introductory chapter has been incorporated which discusses the importance of statistics and why some students find it difficult. One thing has not changed which sets this book apart from others aimed at students: it continues to provide an accessible introduction to the wide range of statistics that are employed by professional researchers. Students using earlier editions of the book will by now often be well into teaching and research careers of their own. We hope that these further enhancements may encourage them to keep *Introduction to Statistics in Psychology* permanently on their desks while they instruct their students how to do statistics properly.

We have considered very carefully the need for instruction into how to compute statistics using SPSS and other computer programs. Our approach in this book is to provide the basic steps needed for the computation but we have added a number of screenshots to help the reader with the analysis. Students of today are very familiar with computers and many do not need overly detailed instructions. Too much detailed step-by-step instruction tends to inhibit exploration of the program – trying things out simply to see what happens and using one's intelligence and a bit of knowledge to work out what things mean. Students can become fixated on the individual steps and fail to learn the complete picture of doing statistics using SPSS or other computer programs. In the end, learning to use a computer program is quicker if the user takes some responsibility for their learning. Much of our daily use of computers in general is on a trial and error basis (we don't need step-by-step instructions to use Facebook or eBay) so why should this be different for statistics programs? How many of us read instructions for the iPhone in detail before trying things out? Of course, there is nothing unusual about tying statistics textbooks to computer packages such as SPSS Statistics. Indeed, our *Introduction to SPSS Statistics in Psychology* is a good example of this approach. It provides just about the speediest and most thorough introduction to doing psychological statistics on SPSS. Unfortunately, SPSS is not the complete answer to the statistical needs of psychologists. It simply does not do everything that students (and professionals for that matter) need to know about. Some of these things are very simple and easily computed by hand if instructions are provided. Other things do require computer programs other than SPSS when procedures are not available on SPSS. We think that ideally psychologists should know the statistics which their discipline needs and not simply those that SPSS provides.

SPSS is very good at what it does but there are times when additional help is needed. This is why we introduce students to other programs which will be helpful to them when necessary. One of the most important features of SPSS Statistics is that it is virtually universally available to students for little or no cost thanks to site licensing agreements. Unfortunately, this is not true of other commercial statistics software. For that reason we have suggested and recommended programs which are essentially free for the user. The Web has a surprisingly large amount of such software to carry out a wide range of statistical routines. A few minutes using Google or some other search engine will often be bountifully productive. Some of these programs are there to be downloaded but others, applets, are instantly available for calculations. We have added at the end of each chapter, advice on the use of software.

This does not mean that we have abandoned responsibility for teaching how statistics works in favour of explaining how to press keys on a computer keyboard. Although we think it best that statistics are computed using statistics programs because the risk of simple calculation errors is reduced, it seems to us that knowing how to go about doing the calculations that computer programs will do for you leads to an understanding of statistics which relying on computers alone does not. So we have included in this edition sections entitled 'Explaining statistics' which are based on hand calculation methods which should help students understand better what the computer program does (more or less) when it is used to do that calculation. Statistical techniques, after all, are little more than the mathematical steps involved in their calculation. Of course, they may be ignored where this level of knowledge is not required.

The basic concept of the book remains the same – a modular statistics package that is accessible throughout to a wide ability range of students. We have attempted to achieve this while being as rigorous as possible where rigour is crucial. Ultimately this is a book for students, though its emphasis on statistics in practice means that it should be valuable to anyone seeking to familiarise themselves with the vast majority of common statistical techniques employed in modern psychology and related disciplines. Not all chapters will be useful to everyone but the book, taken as a whole, provides a sound basis for learning the statistics which professional psychologists use. In this sense, it eases the transition from being a student to being a professional.

# Acknowledgements


## Authors' acknowledgements

We could not have produced this book without the skills and hard work of a number of individuals. Indeed, over the years, many people have contributed in a variety of ways which have helped to make the book what it is. Their contribution is highly valued by us but we would like to mention by name some of those who have been involved in this new edition. In no particular order they are:

Ros Woodward, who was the copy editor for this edition. Her ability to turn the text design brief into the final layout is remarkable. At the same time, she spots so many typos and other problems in the manuscript that we are convinced that she has superhuman powers.

Kevin Ancient supplied the text design without which the book would be far less readable and attractive to look at.

Sue Gard was the proof reader this time. This is a really difficult job for this book as you can imagine. Not only have the words got to be checked but the numbers too. She did a fabulous job of correcting the proofs and checking that we had no gone astray.

Kim Stringer prepared the index. This is a really important job for the user and you will find it all the easier to navigate the book thanks to Kim.

Nicola Woowat designed the cover which probably made you want to pick the book up in the first place. It looks good on your bookshelf thanks to her.

Kerrie Morton and Kay Holman were the production controllers. They do all the liaison work with typesetters and printers and keep things on schedule.

Mary Lince was the project editor. She is therefore a super-efficient master/mistress of the intricacies of publishing who we hold in awe. There is nothing that she can't do.

Janey Webb was the acquisitions editor. Her job is to make our lives unbearable in the nicest possible way. We constantly make changes to improve the manuscript to ensure that her voracious appetite for the best possible manuscripts is satisfied no matter how temporarily. She is a constant strength.

Finally, Jane Lawes has been the editorial assistant on this and other of our books for the past few years. She is leaving to go to pastures new and so this is a timely moment to wish her well in the future. We are grateful for everything that she has contributed and forgive her for getting us to do things that we didn't want to do!

Thanks to everyone for everything, especially their patience with us.

*Dennis Howitt and Duncan Cramer*

## Publisher's acknowledgements

We are grateful to the following for permission to reproduce copyright material:

### Figures

Figure 33.10 from The relation of formal education to ethnic prejudice: its reliability, validity and explanation, *European Journal of Social Psychology*, 25, pp. 41–56, Figure 1, p. 52 (Wagner, U. and Zick, A. 1995). Copyright © 1995 by John Wiley & Sons Ltd. Reproduced with permission of John Wiley & Sons Ltd; Figures 40.6, 40.7, 40.8 from G*Power.

### Screenshots

Screenshots 36.1, 36.2, 36.3, 36.4, 36.5, 36.6 from The Meta-Analysis Calculator, http://www.lyonsmorris.com/lyons/metaAnalysis/index.cfm, reproduced with permission from Larry C. Lyons; Screenshots 40.1, 40.2, 40.3 from G*Power.

SPSS screen images are reprinted courtesy of International Business Machines Corporation, © International Business Machines Corporation. SPSS Inc. was acquired by IBM in October, 2009. IBM, the IBM logo, ibm.com, and SPSS are trademarks or registered trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at 'IBM Copyright and trademark information' at www.ibm.com/legal/copytrade.shtml.

### Tables

Tables Appendix I.1, Appendix I.2 adapted from *Fundamentals of Behavioural Statistics*, McGraw-Hill (Runyon, R.P. and Haber, A. 1989) Table I, The McGraw-Hill Companies, Inc; Significance Table 19.3 adapted and extended from Table I of R.P. Runyon and A. Haber (1989), *Fundamentals of Behavioural Statistics*, McGraw-Hill, The McGraw-Hill Companies, Inc; Table 31.11 adapted from Motivational and informational functions and consquences of children's attention to peers' work, *Journal of Educational Psychology*, 87(3), pp. 347–60 (Butler, R. 1995), published by APA; Table 32.3 adapted from Relationship of gender, self-esteem, social class and racial identity to depression in blacks, *Journal of Black Psychology*, 20(2), pp. 157–74 (Munford, M.B. 1994), Copyright © 1994, Association of Black Psychologists. Reprinted by Permission of Sage Publications; Table 33.2 from The relation of formal education to ethnic prejudice: its reliability, validity and explanation, *European Journal of Social Psychology*, 25, pp. 41–56 (Wagner, U. and Zick, A. 1995), John Wiley & Sons.

### Text

Extract on page 437 after Motivational and informational functions and consequences of children's attention to peers' work, *Journal of Educational Psychology*, 87(3), pp. 347–60, p. 350 (Butler, R. 1995), published by APA; Extract on page 471 from The relation of formal education to ethnic prejudice: its reliability, validity and explanation, *European Journal of Social Psychology*, 25, pp. 53–4 (Wagner, U. and Zick, A. 1995), John Wiley & Sons.

In some instances we have been unable to trace the owners of copyright material, and we would appreciate any information that would enable us to do so.

CHAPTER 1

# Why statistics?

## Overview

- Students do not regard statistics positively, research shows. More importantly, evidence suggests that a poor attitude towards statistics leads to poor learning. Student culture tends to reinforce what is bad in the learning environment for statistics.
- A student's experience within the school environment especially determines their attitudes to mathematics which in its turn impacts on their expectations concerning learning statistics.
- There is a mistaken belief among students that statistics is not central to professional work in psychology and other related careers. Why study something which is unnecessary for psychological work? The truth is quite different. Professional psychologists do use research based on quantitative methods and statistics in their work. Furthermore they are frequently expected to do relevant psychological research as part of their work as psychologists. Many other professions employ statistics routinely and so a good working knowledge of statistics puts psychology students at an advantage in the employment market.
- Learning statistics can be made hard simply because psychologists often employ old and outmoded statistical ideas. Some of these ideas are not only unhelpful but also unworkable. This can only contribute to the fog of confusion surrounding statistics experienced by many students. Textbook writers are frequently guilty of perpetuating these counterproductive ideas.
- Too much emphasis is placed on significance testing. This encourages students to overlook other major contributions of statistics to dealing with the problems inherent in research. It is important to understand the extensive nature and variety of statistics in psychology.
- The mathematical skills needed to develop a good working knowledge of statistics are few in number and well within the capabilities of most students. Even where these have been forgotten, they can be quickly learnt by a motivated student.

## 1.1 Introduction

For many psychology students the formula is simple: statistics = punishment. Statistics is 'sadistics'. Students often find a less palatable subject than statistics unimaginable. The majority would steer well clear of statistics given the choice. All in all, this amounts to a very unpromising learning environment. We usually do best when studying things that we are interested in and want to study. A modern training in psychology inevitably includes statistics – the very thing that students want to avoid. It is not surprising, then, that statistics is a problem area for many students. No two learners are alike, of course, and there is a minority of students who are much more positive towards learning statistics. And we should not forget the poor soul whose job it is to teach statistics to such reluctant students. At best this would appear to be a challenge, at worst an impossibility. Student ratings of statistics modules can bring tears to the eyes of all but the most classroom weary and hardened of lecturers. All round, what could be more unsatisfactory?

Why not just abandon the enterprise and leave statistics out of psychology degrees? What could be more simple? There are many good reasons why this cannot and will not happen. Statistics fills an important and central role in psychology and much psychological research is unthinkable without statistics. Wait a minute – statistics may be essential to many kinds of psychological research but surely there are many psychologists who help people immeasurably but who never do research? In the past this may have been the case but no longer. Most modern psychology careers are fundamentally tied to research in some way. Once this might have meant that psychologists working in fields such as education and mental health merely had to keep up with the relevant published research of others – i.e. the idea of evidence-based practice. Nowadays it is a much more difficult and complex situation. The majority of working psychologists are expected to do research as an aspect of their employment. That is, modern psychologists are practitioner-researchers. As an example, many psychologists working for the forensic prison services contribute much of the research to their particular field of work. Not for a long time has research been purely what academic psychologists do and it is increasingly what every psychologist does. This is also true for many of the other professions that psychology graduates may enter. We are living in an information-based society and research provides a great deal of that information in the modern world. The bottom line of all of this is that basic statistical skills as well as research skills are generally advantageous in the employment market.

### Students and statistics

Unlike most other disciplines, statistics (along with mathematics) is generally negatively evaluated in our culture. The average person in the street probably has an attitude to statistics without knowing anything much about what the discipline involves. That attitude is unlikely to be that statistics is an important, valuable and central part of modern life. Instead, many will groan at the very mention of the word. Hackneyed old phrases such as 'you can prove anything with statistics' and 'lies, damned lies and statistics' will be trotted out to dismiss its achievements. Of course, misleading with statistics is possible but it is not the objective of most statisticians. A few minor adjustments to a graph can lead to a grossly misleading impression at a stroke. A modest growth or decline in a graph may be dramatically changed to seem miraculous or calamitous. But such an important part of modern life as statistics deserves greater respect than this.

The word statistics comes from the Latin for State (as in nation). Statistics originally was the information collected by the State to help governments in their decision-making.

The government's appetite for such figures is prodigious and all of us are affected by them in some way. Pay, pensions and taxes are all partly determined by statistical data as well as where schools and colleges are built. And, of course, we are all part of statistics. Few modern professions do not use statistics in some way. Big supermarkets use it, small charities use it, the health services use it – you name it and they probably use statistics-based research. Without some statistical knowledge, doing and understanding research is very difficult and a precarious occupation.

Nevertheless, on a personal level, students study psychology to study psychology – not to study statistics. Superficially it is possible to study psychology without statistics. Get deeper into psychology and some knowledge of statistics becomes increasingly necessary. This is not to deny the growing interest in qualitative research which does not involve statistics almost by definition. Much valuable research is done by qualitative researchers (Howitt, 2013). But this does not mean that quantitative statistical methods have released their grip on psychological research to any significant extent. Both qualitative and quantitative research seem to be prospering in psychology. Statistics and psychology are seemingly forever intertwined. OK, we are not serious that statistics is taught just to punish students – no matter that sometimes it may feel that way. You might try an alternative view of statistics – that it is a sort of cuddly friend which will help you in all sorts of ways. We are serious here. Criticisms of the dominance of statistics in psychology are common, of course. As much as anyone else, we are as against the mindless application of statistics in psychology for its own sake. Psychology may seem obsessed with a few limited statistical topics such as significance testing but this is to overlook the myriad of more far-reaching positive benefits to be gained from the proper application of modern statistical ideas. Statistics provides a means of finding order in otherwise vast sets of confusing data. Some of this variety of use is illustrated in Figure 1.1.

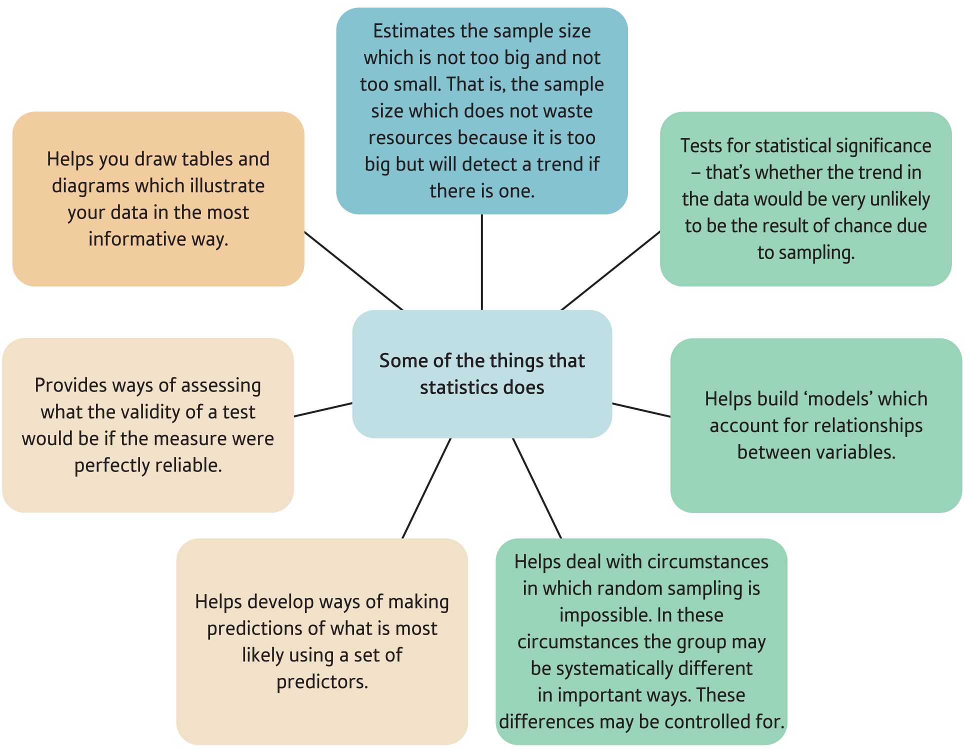



**FIGURE 1.1** Some things that statistics can do for the researcher

## 1.2 Research on learning statistics

Not surprising given the culturally negative view of statistics, the research on psychology students and statistics makes generally depressing reading. The response of student cultures to statistics can just about be summed up with the words trepidation and anxiety. For example, Gordon (2004) surveyed a large number of Australian students about their experience of statistics on psychology courses. Three-quarters said that they would not study statistics but for the fact that it was compulsory. Predominantly they saw it as boring and difficult. These unwilling students felt that statistics was not necessary to psychology or to being a psychologist. They approach statistics as if it were merely a few mechanical procedures that one applies without needing to understand why. One student put it this way to Gordon (1995):

> I have a very pragmatic approach to university, I give them what they want . . . I really do like knowledge for knowledge's sake, but my main motivation is to pass the course.

Although some students try to master the methods and concepts of statistics, they may have difficulty in understanding the importance of statistics. Those who saw statistics as being more personally meaningful in their studies would say things like 'It would probably be useful in whatever job I do' (Gordon, 1995). As might be expected, these more positively orientated students performed a little better in their statistics tests and examinations than the more negative group. The latter were not generally less able students since they did just as well as any other students in their other psychology courses. But not seeing the point of statistics did have a negative impact on their studies. Figure 1.2 provides a broad classification of students in terms of how they see the relevance of statistics and their personal assessment of the discipline.

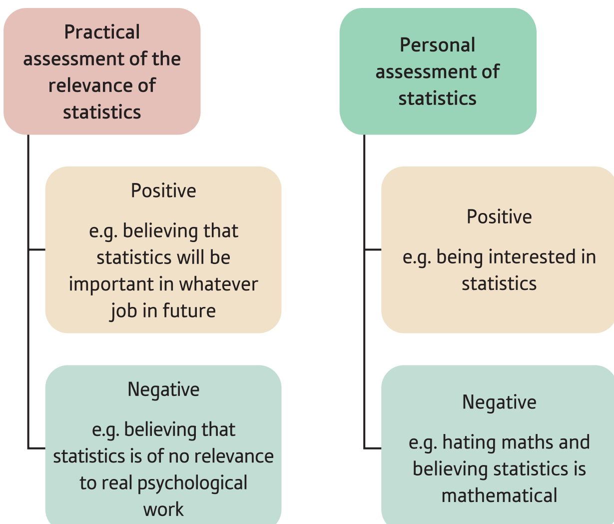



FIGURE 1.2 The responses of students to statistics according to Gordon (1995)

## 1.3 What makes learning statistics difficult?

University staff commonly recognise that teaching statistics involves dealing with problems such as anxieties, beliefs and negative attitudes concerning the subject (Schau, 2003). Indeed, these background issues may be the most important things in the learning process and consequently have a bearing on statistics teaching. University can be an experience full of emotion, and emotion affects learning. This is perhaps more true for a topic such as statistics. Real tears are shed. One student told Gordon (1995), 'I was drowning in statistics' – words which are both emotive and extreme, of course. Being at university and studying statistics follows a long period of personal development through schooling (and for some at work). This background provides the individual with ways of perceiving their own personal learning processes and their education more generally. What they think they know about themselves (e.g. 'I'm no good at maths' or 'I'm an arty sort of person') impacts on their response to statistics. Personal histories, personal experiences, personal needs and personal goals are reflected in their strategies for coping with statistics (Gordon, 2004).

In other words, students bring to learning statistics baggage which may seriously interfere with its learning. Inevitably, high on the list of background factors is one's personal experience of mathematics. There is a strong belief that a high level of mathematical ability is crucial to the learning of statistics. This is reinforced by those universities which require good mathematical qualifications for admission to psychology degrees. Some students may (incorrectly) assume that statistics is beyond their mathematical ability. With so many other demands on their time at university, instead of getting down to studying statistics they may adopt avoidance tactics such as skipping lectures. Furthermore, every statistics class has its own culture in which students influence each other in terms of attitudes to learning statistics. A class dominated by students antagonistic to statistics is not a good learning environment, for example. The problem is that many chosen responses to statistics such as acting silly, talking in statistics lectures or plagiarising the work of other students just do not help. However, the importance of mathematical ability in using statistics effectively is questioned by many, including ourselves, as we shall see.

### But I've always struggled with maths . . .

Research strongly indicates that three factors – anxiety, attitudes and ability (see Figure 1.3) are involved in learning statistics and other somewhat unpopular activities such as learning second languages (Lalonde and Gardner, 1993). A negative attitude towards statistics is associated with poorer performances in statistics to some extent but the other factors are at least equally important. Anxiety plays its part primarily through a specific form of anxiety known as mathematics (math) anxiety. This is more important than trait

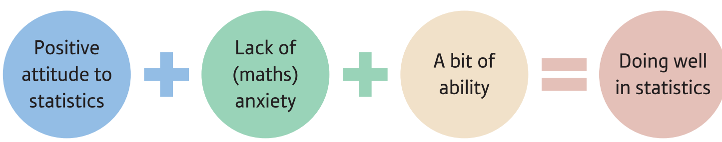



**FIGURE 1.3** The formula for doing well in statistics based on research findings

or general anxiety such as where someone has a generally anxious personality in all sorts of situations. Mathematics anxiety is common among psychology students. Those with higher levels of mathematics anxiety tend to do worst in statistics. To be sure, mathematical ability is associated with better test and examination results, but not to a major extent. Poor mathematical ability has its influence largely because it is associated with increased levels of mathematical anxiety. It is because poorer maths ability leads to increased levels of mathematical anxiety that mathematical anxiety leads to poor learning strategies.

But is statistics particularly mathematical and, if it is, then does it need to be beyond a few basics? Along with others, we would argue that the level of mathematical ability needed to cope with the mathematical part of statistics is not great – fairly minimal in fact. We can safely lay aside the issue of the mathematical ability required to carry out statistical calculations as there are many computer programs such as SPSS and numerous applets on the Web which will do the calculation for you. Indeed, there is not a lot of sense in doing statistical calculations by hand as this invites errors to creep in. Computer programs, so long as you enter the data properly and tell them to do the right thing, will do the calculation without error. However, we do not believe that it is possible to learn statistics without using a little bit of mathematics. Equally, it is not necessary to go into all of the mathematical detail behind a statistical technique in order to understand the reasons why the technique was developed and how it can be used. You will find statistical textbooks for psychologists which fall at these extremes. The idea of statistics without maths or statistics without tears, even, cannot provide the necessary understanding in our view because some of the language of statistics is mathematical in nature. At the same time, books that rejoice in the mathematical intricacies of statistical techniques will lose many of their readers who simply do not have mathematical skills at this level. Best-selling statistics textbooks which appear to be student friendly and full of jokes will sometimes go into the most arcane detail about statistical techniques that are way beyond most of us. This seems to us just as unhelpful as not including any mathematics at all.

Just what mathematical knowledge does one need to get a working insight into statistics? By and large if you understand the concepts of addition, subtraction, multiplication and division then you have the basics. You may get the answers wrong – the question is, do you understand what you are doing? What might you need beyond this? Little more than the following we would say:

- You need to understand the concept of squaring (that is multiplying a number by itself).
- You need to understand the concept of square root (the square root of a number is that number which when multiplied by itself gives the original number).
- It is good too if you understand negative numbers – such as that when multiplying two negative numbers you get a positive number but when you multiply a positive number by a negative number then the result is a negative number. A few minutes trying out positive and negative calculations on a calculator is a good way to refresh yourself of these basics.
- It is preferable if you understand the underlying principles or 'rules' governing mathematical formulae as these are used in statistical formulae but if you don't, your computer does.

Not much else is necessary – if you know what a logarithm is then you are in the ultra-advanced class. So we think that the amount of mathematics needed to make a good statistics student and a skilled user of statistical techniques in research is fairly minimal. Anything that has been forgotten or never learnt will be quickly picked up by a motivated student. Not all lecturers will share this opinion but the overwhelming majority know that students can struggle with statistics and try to provide teaching which serves the needs of all students taking the psychology programme and not the maths-able elite.

If more research evidence is needed, using a formal measure known as the Survey of Attitudes Toward Statistics, Zimprich (2012) was able to show that these attitudes towards statistics are made up of four components:

- *Affect* How positive or negative a student is about statistics (e.g. 'I will like statistics').
- *Cognitive competence* A student's beliefs about their ability and competence to do statistics (e.g. 'I will make a lot of maths errors in statistics').
- *Value* Attitudes concerning the relevance and usefulness of statistics (e.g. 'I use statistics in my everyday life').
- *Difficulty* The student's views about how difficult or easy statistics is (e.g. 'Statistics is a complicated subject').

All of these components were interrelated, as one might expect. When these attitudes were correlated with actual performance in statistics it was clear that attitudes were much more important than actual maths ability in students' performances in statistics. In other words, how a student feels about statistics has a far more tangible effect on their performance on statistical tests and examinations than their mathematical ability.

Irrespective of how mathematical statistics is or isn't, it has to be acknowledged that statistics is a unique and distinctive way of thinking (Ben-Zvi & Garfield, 2004; Ruggeri, Dempster & Hanna, 2011). It is much like mathematics in employing a distinctive language and concepts. Nevertheless it is wrong to think that this statistical language and these concepts have much in common with mathematics. This means that statistics will always be a somewhat 'different' subject irrespective of the curriculum involved. Crucially, statistics is about the use of quantitative research skills in the attempt to answer real research problems. Without being skilful in quantitative research methods, statistics can only partially be understood – and might seem pointless as a consequence. Although research skills take a lot of time and effort to learn, they are very little to do with mathematics – they are primarily about thinking logically. Statistics interfaces with this understanding of research methods in a way which is not simply remembering and then regurgitating a few statistical formulae and ideas when required to do so.

## 1.4 Positive about statistics

So how does one go about having a more positive attitude towards statistics? The answer lies in having an appreciation of what statistics does prior to being exposed to the nitty-gritty or detail taught in the stats lecture room. Take, for example, what is probably the best known statistical research – the national census. We discuss this in Chapter 2. This census, basically, is a questionnaire about all sorts of things of interest to the government and its decision-making, though probably less interesting to the rest of us. The head of every household is required to complete this detailed questionnaire for a particular day usually once every ten years. In the UK this has been going on for over 200 years. It is hard not to think, when the census envelope arrives, 'what a waste of time' and then 'what a waste of money'. This is possibly because we are all aware that researchers use samples. If research always was so comprehensive as to include everyone then little research would ever get done because of the time and expense involved. This is obvious, but only from the hindsight that comes with living in modern times – people had to invent sampling to replace censuses. And this in statistics had its origins in the work of William Gossett.

One of the most famous statistical techniques to impact psychology is the *t*-test (see Chapters 13 and 14) or the Student *t*-test as it is also known. Student was the pen name of William Gosset who had studied chemistry and mathematics at university. He was employed by the Guinness Brewery in Dublin as a 'bright young thing' in the 1890s. Even then, the firm believed in bringing new ideas to the company, thus keeping it abreast with developments. One issue relevant was that of quality control. There are obvious practical problems if every bottle or barrel of beer had to be tested, for example, in order to see if the alcoholic strength was constant throughout all batches. What Gosset did was to work out mathematically a way of estimating the extent that one is likely to be wrong (risks being wrong) if one took samples rather than tested the entire output. By how much are you likely to be wrong (or in error) if you simply took a sample, say, of ten bottles of beer? Of course, you will never know from a sample exactly what the error will be but Gosset was able to estimate what the likely extent of error will be. Put into a formula, this is the idea of standard error which plagues many students on introductory statistics courses. By developing this, Gosset had laid the systematic basis for doing research on samples rather than on everything. Think about it: if it had not been for Gosset's innovation then you would spend your lifetime carrying out your first research study simply because you need to test everyone or everything (the population). So rather than considering William Gosset as some sort of alien, it would be best to regard him as one of the statistical cuddly friends we mentioned earlier!

## Is it statistically significant?

The point of Gosset's revolutionary ideas is probably easy to see when explained in this way. But instead students are introduced to what to them are rather complex formulae and the question 'Are your findings statistically significant?' The question 'Is it significant?' is one of the fixations of psychologists – the question probably sounds like a mantra to students when they first begin to study psychology. So intrusive is the question that for most students, statistics in psychology is about knowing what test of statistical significance to apply in what setting. But this is only a small part of statistics, which provides a whole range of tools to help researchers (and students) address the practical problems of research. Research data can be very simple but also very complex. Statistics helps sort out the complexity and uncertainty involved in understanding your data. Testing for statistical significance merely means assessing whether the trend in your data could have been obtained by choosing a random sample if, in reality, there was no trend in general. That is, how likely is it that the trend could simply be the result of a fortuitous selection of a sample in which there appears to be a trend? (A trend might be, say, athletes scoring more highly on a measure of personal ambition than non-athletes or a relationship between a measure of ability to speak foreign languages and a measure of sociability.)

## What sample size do I need?

Testing for significance needs to be put into context. Really you want to know if there is any support for the ideas underlying your research question and the extent to which the trends in your data are big, little or non-existent. So if we put on our thinking head, and not our 'Is it significant?' head, we would ask rather more sophisticated questions. One would be whether if there really is a trend in our data, i.e. have we got a sample size big enough to show statistical significance for that trend? Statistics can help us with that question by helping us to decide the minimum sample size to show that trend to be statistically significant if there is a trend of a given size in reality rather than just in our data. There would be something perverse about planning research which involved a sample size so small that our findings could never be statistically significant. But that is done all of the time simply because researchers (especially students) do not address the question of minimum sample size properly. Often the advice is given to those asking what sample size to use is that they should use as big a sample size as possible. What does this mean? Possibly it means the largest sample size that you have the resources to collect. But the availability of resources is hardly a satisfactory basis on which to formulate research – that would be a bit like going shopping with the objective of spending money for its own sake rather than to buy something that is necessary. For socially important research, funding may be fairly readily available such as in the case of a cure for cancer. Does this mean that all resources should be put into a particular research project? Not really, as this might well be a complete waste of money when the research question could be addressed satisfactorily with a fairly small sample size.

Research takes a lot of time, effort and organisation. So naturally many students will ask the perfectly reasonable question 'What sample size do I need?', but frequently they will fail to get a satisfactory answer. This is partly because too many psychologists regard 'statistical significance' as the be all and end all of research. The question that the student is asking is actually far more sophisticated than the answers they receive. The consequence of telling a student that they should get the biggest sample they can or that they should have a minimum sample of 50 or 100 or whatever is bewilderment on the part of the student, who realises but can't explain why these answers are inadequate. Statistics is about sophisticated decision-making concerning what can be said on the basis of the research but also about whether to proceed further with a particular line of inquiry. Statistical significance has a part to play in this decision-making but it does not mean that research findings are significant in any other respect – they may be uninteresting, they may not be of any practical significance, and they may not address any theoretically important issues yet they are deemed statistically significant. It is far better if students understand that there are many issues that a researcher needs to address in their work way beyond statistical significance – while accepting that statistical significance is important in its own way. Many chapters in this book (such as Chapters 11 and 18) discuss statistical significance but the important question of sample size is addressed only in Chapter 40.

## Is there a trend in my data?

What the student really wants to know is the optimum sample size if there is truly a trend in the data (rather than one that is the consequence of the vicissitudes of sampling). Just taking the largest sample possible may result in a sample that is far too small or far too large. Both of these are unsatisfactory. A too-small sample might mean that your data do not reach statistical significance even where there is really in fact a trend in the real world. This research would be a waste of money and other resources as it cannot answer the question asked satisfactorily. A too-large sample might mean that very small and uninteresting trends in the data are statistically significant. Even where there is a substantial trend in the data, the too-large sample will nevertheless waste time and other resources because the question asked can be satisfactorily answered with a rather smaller sample. Imagine a big medical trial. This is likely to be expensive and every extra person in the research sample costs a great deal of money. This may be money wasted unnecessarily. For this reason, organisations that finance medical research expect the researcher to be able to say just what sample size is big enough to reach statistical significance if there is a trend in reality but not so big that a small, uninteresting trend is detected. What makes for an interesting trend is one which is sufficiently large that it has economic, commercial or some other form of potential. The size of the interesting trend depends on