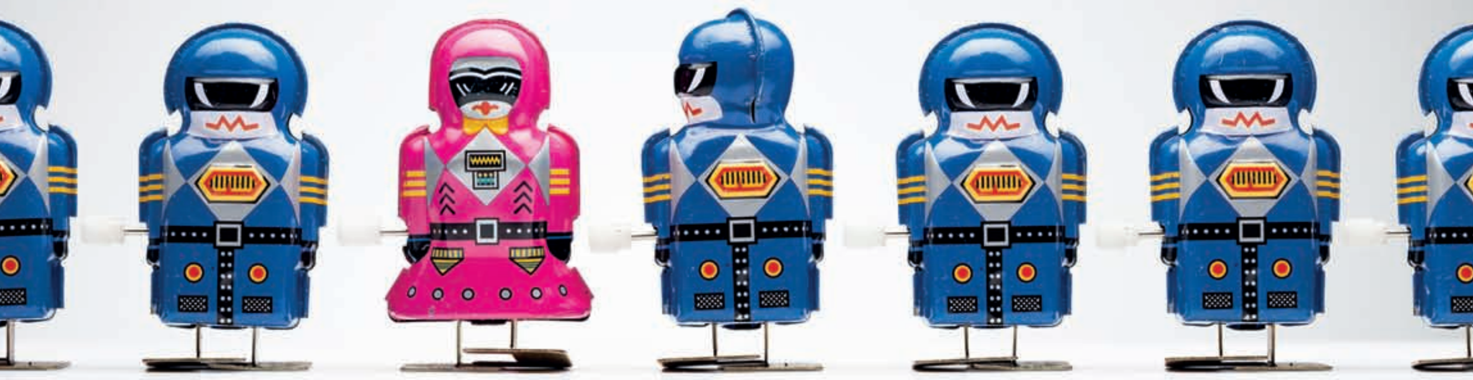
Understanding Statistics in Psychology with SPSS

Dennis Howitt and Duncan Cramer

Seventh Edition



Pearson

Understanding Statistics in Psychology with SPSS

Understanding Statistics in Psychology with SPSS

Seventh 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 2014 (print and electronic)
Seventh edition published 2017 (print and electronic)

ISBN: 978-1-292-13421-5 (print)
978-1-292-13424-6 (PDF)
978-1-292-13425-3 (ePub)

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
Names: Howitt, Dennis, author. | Cramer, Duncan, 1948- author.
Title: Understanding statistics in psychology with SPSS / Dennis Howitt, Loughborough University, Duncan Cramer, Loughborough University.
Other titles: Introduction to statistics in psychology.
Description: Seventh Edition. | New York : Pearson, 2017. | Revised edition of the authors' Introduction to statistics in psychology, 2013.
Identifiers: LCCN 2016047666 | ISBN 9781292134215 (print) | ISBN 9781292134246 (pdf) | ISBN 9781292134253 (epub)
Subjects: LCSH: Psychometrics.
Classification: LCC BF39 .H74 2017 | DDC 150.1/5195–dc23
LC record available at https://lccn.loc.gov/2016047666

10 9 8 7 6 5 4 3 2 1
21 20 19 18 17

Print edition typeset in 9.5/12 pt Sabon LT Pro by Spi Global
Printed in Slovakia by Neografia

NOTE THAT ANY PAGE CROSS REFERENCES REFER TO THE PRINT EDITION

Brief contents

Contents

Companion Website

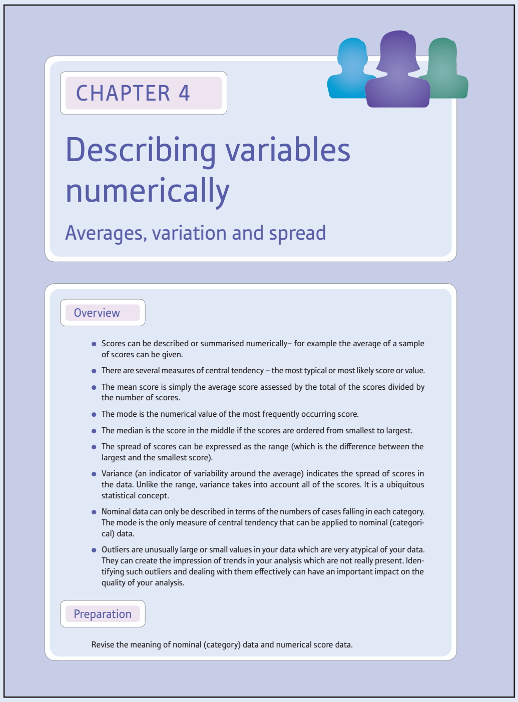
For open-access **student resources** specifically written to complement this textbook and support your learning, please visit **www.pearsoned.co.uk/howitt**

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

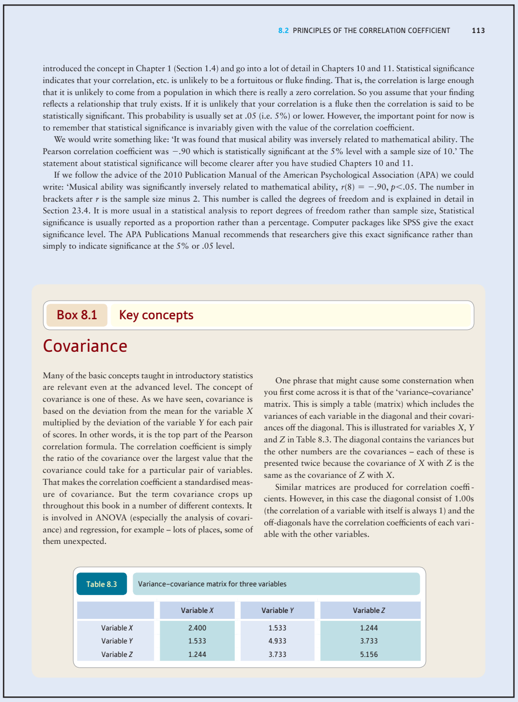
Averages, variation and spread

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 or value.
- 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 (categorical) 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 analysis.

Preparation

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



8.2 PRINCIPLES OF THE CORRELATION COEFFICIENT 113

introduced the concept in Chapter 1 (Section 1.4) and go into a lot of detail in Chapters 10 and 11. Statistical significance indicates that your correlation, etc. is unlikely to be a fortuitous or fluke finding. That is, the correlation is large enough that it is unlikely to come from a population in which there is really a zero correlation. So you assume that your finding reflects a relationship that truly exists. If it is unlikely that your correlation is a fluke then the correlation is said to be statistically significant. This probability is usually set at .05 (i.e. 5%) or lower. However, the important point for now is to remember that statistical significance is invariably given with the value of the correlation coefficient.

We would write something like: 'It was found that musical ability was inversely related to mathematical ability. The Pearson correlation coefficient was −.90 which is statistically significant at the 5% level with a sample size of 10.' The statement about statistical significance will become clearer after you have studied Chapters 10 and 11.

If we follow 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 detail in Section 23.4. It is more usual in a statistical analysis to report degrees of freedom rather than sample size, Statistical significance is usually reported as a proportion rather than a percentage. Computer packages like SPSS give the exact significance level. The APA Publications Manual recommends that researchers give this exact significance rather than simply to indicate significance at the 5% or .05 level.

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 based on the deviation from the mean for the variable X multiplied by the deviation of the variable Y for each pair of scores. 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. That makes the correlation coefficient 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 when you first come across it is that of the 'variance–covariance' matrix. 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 X.

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 variables.

Table 8.3 Variance–covariance matrix for three variables

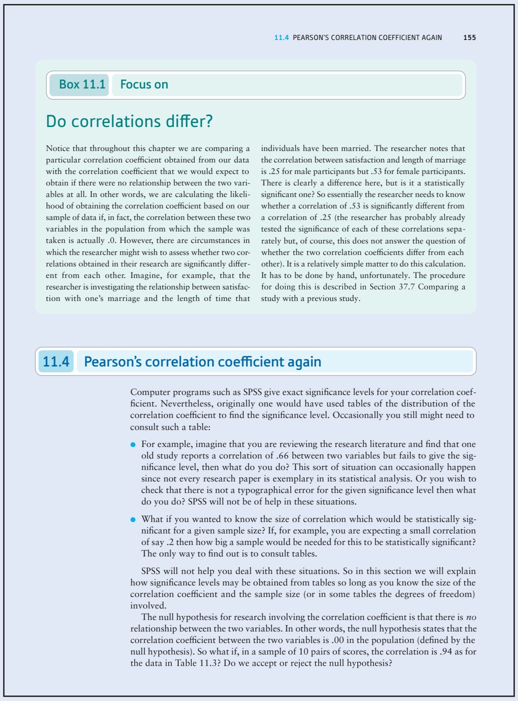
	Variable *X*	Variable *Y*	Variable *Z*
Variable *X*	2.400	1.533	1.244
Variable *Y*	1.533	4.933	3.733
Variable *Z*	1.244	3.733	5.156

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.



11.4 PEARSON'S CORRELATION COEFFICIENT AGAIN 155

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. 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 .25 for male participants but .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 .53 is significantly different from a correlation of .25 (the researcher has probably already tested the significance of each of these correlations separately 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 37.7 Comparing a study with a previous study.

11.4 Pearson's correlation coefficient again

Computer programs such as SPSS give exact significance levels for your correlation coefficient. Nevertheless, originally one would have used tables of the distribution of the correlation coefficient to find the significance level. Occasionally you still might need to consult such a table:

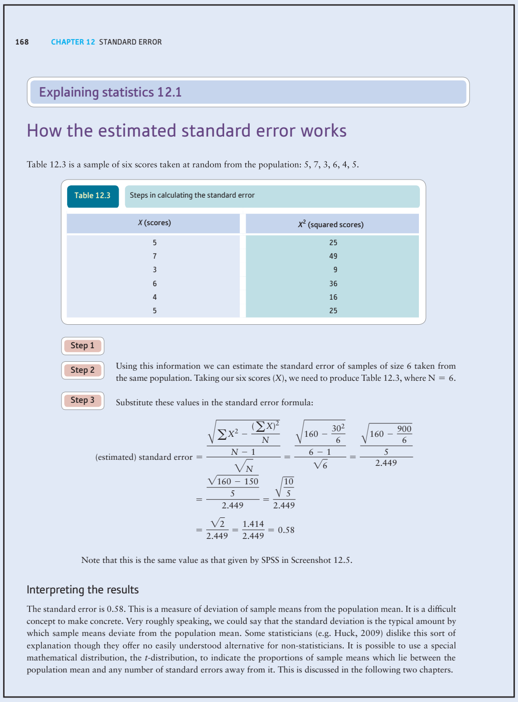
- For example, imagine that you are reviewing the research literature and find that one old study reports a correlation of .66 between two variables but fails to give the significance level, then what do you do? This sort of situation can occasionally happen since not every research paper is exemplary in its statistical analysis. Or you wish to check that there is not a typographical error for the given significance level then what do you do? SPSS will not be of help in these situations.
- What if you wanted to know the size of correlation which would be statistically significant for a given sample size? If, for example, you are expecting a small correlation of say .2 then how big a sample would be needed for this to be statistically significant? The only way to find out is to consult tables.

SPSS will not help you deal with these situations. So in 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 in some tables the 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 states that the correlation coefficient between the two variables is .00 in the population (defined by the null hypothesis). So what if, in a sample of 10 pairs of scores, the correlation is .94 as for the data in Table 11.3? Do we accept or reject the null hypothesis?

Focus on

Explore particular concepts in more detail.



168 CHAPTER 12 STANDARD ERROR

Explaining statistics 12.1

How the estimated standard error works

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

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

Step 1

Step 2 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 3 Substitute these values in the standard error formula:

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

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

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

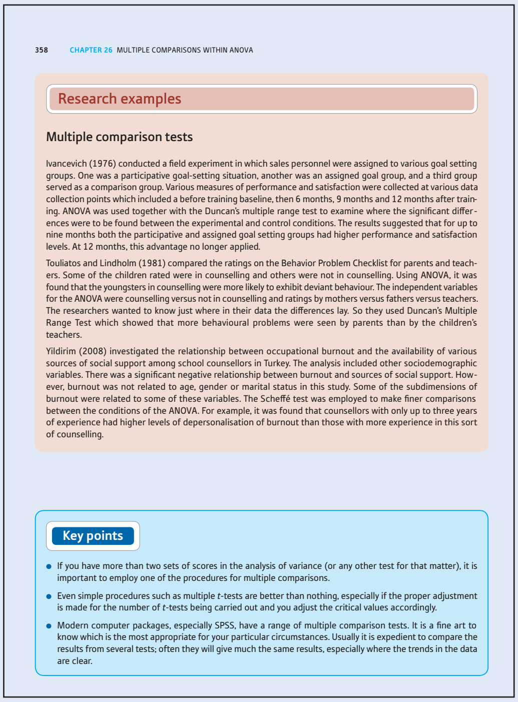
Note that this is the same value as that given by SPSS in Screenshot 12.5.

Interpreting the results

The standard error is 0.58. This is a measure of deviation of sample means from the population mean. It is a difficult concept to make concrete. Very roughly speaking, we could say that the standard deviation is the typical amount by which sample means deviate from the population mean. Some statisticians (e.g. Huck, 2009) dislike this sort of explanation though they offer no easily understood alternative for non-statisticians. It is possible to use a special mathematical distribution, the t-distribution, to indicate the proportions of sample means which lie between the population mean and any number of standard errors away from it. This is discussed in the following two chapters.

Explaining statistics

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



358 CHAPTER 26 MULTIPLE COMPARISONS WITHIN ANOVA

Research examples

Multiple comparison tests

Ivancevich (1976) conducted a field experiment in which sales personnel were assigned to various goal setting groups. One was a participative goal-setting situation, another was an assigned goal group, and a third group served as a comparison group. Various measures of performance and satisfaction were collected at various data collection points which included a before training baseline, then 6 months, 9 months and 12 months after training. ANOVA was used together with the Duncan's multiple range test to examine where the significant differences were to be found between the experimental and control conditions. The results suggested that for up to nine months both the participative and assigned goal setting groups had higher performance and satisfaction levels. At 12 months, this advantage no longer applied.

Touliatos and Lindholm (1981) compared the ratings on the Behavior Problem Checklist for parents and teachers. Some of the children rated were in counselling and others were not in counselling. Using ANOVA, it was found that the youngsters in counselling were more likely to exhibit deviant behaviour. The independent variables for the ANOVA were counselling versus not in counselling and ratings by mothers versus fathers versus teachers. The researchers wanted to know just where in their data the differences lay. So they used Duncan's Multiple Range Test which showed that more behavioural problems were seen by parents than by the children's teachers.

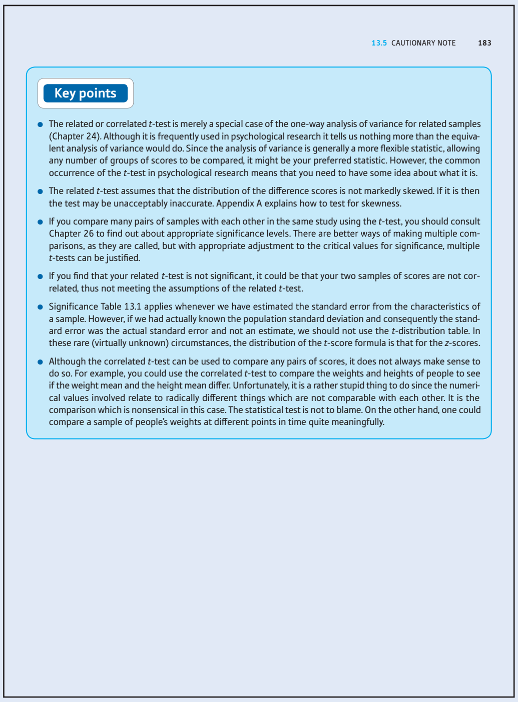
Yildirim (2008) investigated the relationship between occupational burnout and the availability of various sources of social support among school counsellors in Turkey. The analysis included other sociodemographic variables. There was a significant negative relationship between burnout and sources of social support. However, burnout was not related to age, gender or marital status in this study. Some of the subdimensions of burnout were related to some of these variables. The Scheffé test was employed to make finer comparisons between the conditions of the ANOVA. For example, it was found that counsellors with only up to three years of experience had higher levels of depersonalisation of burnout than those with more experience in this sort of counselling.

Key points

- If you have more than two sets of scores in the analysis of variance (or any other test for that matter), it is important to employ one of the procedures for multiple comparisons.
- Even simple procedures such as multiple *t*-tests are better than nothing, especially if the proper adjustment is made for the number of *t*-tests being carried out and you adjust the critical values accordingly.
- Modern computer packages, especially SPSS, have a range of multiple comparison tests. It is a fine art to know which is the most appropriate for your particular circumstances. Usually it is expedient to compare the results from several tests; often they will give much the same results, especially where the trends in the data are clear.

Research examples

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



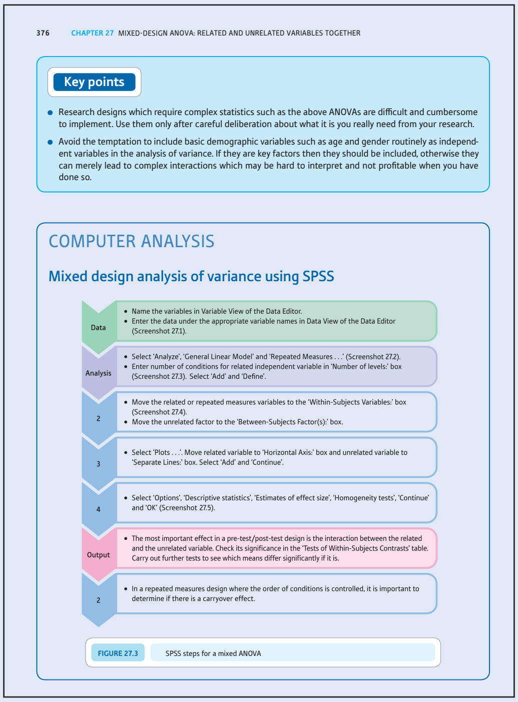
13.5 CAUTIONARY NOTE 183

Key points

- The related or correlated *t*-test is merely a special case of the one-way analysis of variance for related samples (Chapter 24). 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 26 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.


376 CHAPTER 27 MIXED-DESIGN ANOVA: RELATED AND UNRELATED VARIABLES TOGETHER

Key points

- Research designs which require complex statistics such as the above ANOVAs are difficult and cumbersome to implement. Use them only after careful deliberation about what it is you really need from your research.
- Avoid the temptation to include basic demographic variables such as age and gender routinely as independent variables in the analysis of variance. If they are key factors then they should be included, otherwise they can merely lead to complex interactions which may be hard to interpret and not profitable when you have done so.

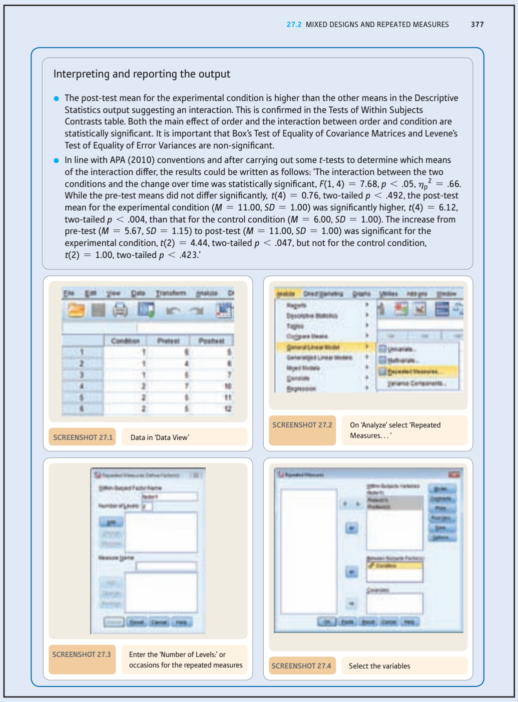
COMPUTER ANALYSIS

Mixed design analysis of variance using SPSS

FIGURE 27.3 SPSS steps for a mixed ANOVA

Computer analysis

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


27.2 MIXED DESIGNS AND REPEATED MEASURES 377

Interpreting and reporting the output

- The post-test mean for the experimental condition is higher than the other means in the Descriptive Statistics output suggesting an interaction. This is confirmed in the Tests of Within Subjects Contrasts table. Both the main effect of order and the interaction between order and condition are statistically significant. It is important that Box's Test of Equality of Covariance Matrices and Levene's Test of Equality of Error Variances are non-significant.
- In line with APA (2010) conventions and after carrying out some *t*-tests to determine which means of the interaction differ, the results could be written as follows: 'The interaction between the two conditions and the change over time was statistically significant, $F(1, 4) = 7.68, p < .05, \eta_p^2 = .66$. While the pre-test means did not differ significantly, $t(4) = 0.76$, two-tailed $p < .492$, the post-test mean for the experimental condition ($M = 11.00, SD = 1.00$) was significantly higher, $t(4) = 6.12$, two-tailed $p < .004$, than that for the control condition ($M = 6.00, SD = 1.00$). The increase from pre-test ($M = 5.67, SD = 1.15$) to post-test ($M = 11.00, SD = 1.00$) was significant for the experimental condition, $t(2) = 4.44$, two-tailed $p < .047$, but not for the control condition, $t(2) = 1.00$, two-tailed $p < .423$.'

SCREENSHOT 27.1 Data in 'Data View'

SCREENSHOT 27.2 On 'Analyze' select 'Repeated Measures . . .'

SCREENSHOT 27.3 Enter the 'Number of Levels:' or occasions for the repeated measures

SCREENSHOT 27.4 Select the variables

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 seventh edition of what has been retitled *Understanding Statistics in Psychology with SPSS* will contribute even more to the student learning experience. A number of changes have been made to this end. One thing that 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 *Statistics in psychology using SPSS* permanently on their desks while they instruct their students how to do statistics properly. The abbreviation SPSS initially stood for Statistical Software for the Social Sciences. Although the official name of the latest release at the time of publication is IBM SPSS Statistics 23.0 we shall refer to it throughout this book as SPSS because it is shorter, most users refer to it this way and the first letter of the original acronym actually refers to Statistical and so to add Statistics again seems repetitive. For most users of SPSS, SPSS versions have changed little since SPSS 13 came out in 2005, so this book will also be suitable for those using these earlier releases.

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. Indeed, our *Introduction to SPSS 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. 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 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 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

Producing a statistics book requires a tremendous effort from those who turn the manuscript into the book which you have in your hands. They deserve the lion's share of the credit for making the book as good as it is. It is all too easy with a topic like statistics to introduce errors which put the reader off even though they are simple typos. So we are especially grateful to everyone involved for their professionalism, talent and niceness.

The commissioning editor is probably the person that we have the most contact with and we have been blessed with working with Janey Webb (Publisher) at Pearson not just for several editions of this book but several other books too. She has a complicated role which includes getting the best out of the authors. This she achieves in the most charming, gentle manner that any author would just succumb and obey! Janey was on maternity leave at the time of the initial planning of this edition and Lina Aboujieb took over in Janey's place. We enjoyed working with her greatly.

Turning to the production side, Jennifer Sargunar (Managing Producer) was in charge of turning the manuscript into a book. She organised everything brilliantly making the whole process seem organised, structured and smooth running. We cannot help but be impressed and very grateful that she worked on the book with us. Maggie Harding provided the great cover design and Kevin Ancient the text design. You will have hardly failed to realise the complexity of the material and it is Kevin's design which structures confusion into something comprehensible.

Particularly important to authors are the copy editor and proof reader. This is especially the case with this book where the smallest typo might throw confusion into the works and deter the reader. Ros Woodward was the copy editor for this book. Her main role was to apply the text design to the manuscript as well as being the first-run proof reader. As always, she did a fantastic job. Getting the text design right makes a complicated book much easier to read and study. Jen Halford was the proof reader. What can we say? What superlatives have we not used? Jen was amazing at not just spotting the usual typos and layout problems but finding computational mistakes too. It is not possible to overestimate the quality of Jen's work.

Finally we would like to thank the following academic reviewers for their valuable input:

Dr Emma O'Dwyer, Kingston University

Dr Fraser Milton, University of Exter

Dr Philip Ansell, Newcastle University

Dr Helen Keyes, Anglia Ruskin University

Dennis Howitt, Duncan Cramer

Publisher's acknowledgements

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

Figures

Figure 35.10 after 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), © 1995 by John Wiley & Sons, Ltd.; Figures 40.6, 40.7, 40.8, 40.9 from G*Power, © Copyright 2010–2016 Heinrich-Heine-Universität Düsseldorf.

Screenshots

Screenshots 1.1, 1.2, 1.3, 2.1, 2.2, Appendix B2.2, 2.3, Appendix B2.3, 2.4, Appendix B2.4, 2.5, Appendix B2.5, Appendix B2.6, 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 7.1, 7.2, 7.3, 7.4, 7.5, 7.6, 8.1, 8.2, 8.3, 8.4, 8.5, 8.6, 8.7, 8.8, 8.9, 8.10, 8.11, 8.12, 9.1, 9.2, 9.3, 9.4, 9.5, 10.1, 10.2, 10.3, 10.4, 10.5, 11.1, 11.2, 11.3, 11.4, 11.5, 11.6, 12.1, 12.2, 12.3, 12.4, 12.5, 13.1, 13.2, 13.3, 13.4, 13.5, 14.1, 14.2, 14.3, 14.4, 14.5, 16.1, 16.2, 16.3, 16.4, 16.5, 16.6, 18.1, 18.2, 18.3, 18.4, 18.5, 21.1, 21.2, 21.2, 21.3, 21.4, 21.5, 21.6, 22.1, 22.2, 22.3, 22.4, 22.5, 23.1, 23.2, 23.3, 23.4, 23.5, 24.1, 24.2, 24.3, 24.4, 24.5, 25.1, 25.2, 25.3, 25.4, 25.5, 26.1, 26.2, 26.3, 26.4, 26.5, 26.6, 27.1, 27.2, 27.3, 27.4, 27.5, 27.6, 28.1, 28.2, 28.3, 28.4, 28.5, 28.6, 28.7, 28.8, 28.9, 29.1, 29.2, 29.3, 29.4, 29.5, 30.1, 30.2, 30.3, 30.4, 30.5, 31.1, 31.2, 31.3, 31.4, 31.5, 31.6, 32.1, 32.2, 32.3, 32.4, 32.5, 33.1, 33.2, 33.3, 33.4, 34.1, 34.2, 34.3, 34.4, 34.5, 35.1, 35.2, 35.3, 35.4, 35.5, 36.1, 36.2, 36.3, 36.4, 36.5, 36.6, 38.1, 38.2, 38.3, 38.4, 38.5, 39.1, 39.2, 39.3, 39.4, 39.5, 41.1, 41.2, 41.3, 42.1, 42.2, 42.3, 42.4, 42.5, 43.1, 43.3, 43.4, 43.5 from SPSS Statistics screenshot image International Business Machines Corporation, screenshots reprinted courtesy of International Business Machines Corporation, © International Business Machines Corporation. SPSS was acquired by IBM in October, 2009. IBM, the IBM logo, ibm.com, and SPSS are trademarks of International Business Machines Corp., 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.; Screenshots 37.1, 37.2, 37.3, 37.4, 37.5, 37.6 from The Meta-Analysis Calculator, http://www.lyonsmorris.com/lyons/metaAnalysis/index.cfm, Reproduced by permission of Larry C.Lyons; Screenshots 40.1, 40.2, 40.3 from G*Power, © Copyright 2010–2016 Heinrich-Heine-Universität Düsseldorf.

Tables

Table 34.3 after Relationship of Gender, Self-Esteem, Social Class, and Racial Identity to Depression in Blacks, *Journal of Black Psychology* Vol 20 (2), pp. 157–174 (Munford, Maria B), Copyright © 1994 by Association of Black Psychologists; Table 35.2 after 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), © 1995 by John Wiley & Sons, Ltd.

Text

Extract 35.4 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), © 1995 by John Wiley & Sons, Ltd.

CHAPTER 1

Why statistics?

Overview

- Students do not generally approach learning statistics positively. Everyone knows this but it is demonstrated by research too. More importantly, this 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 that is unnecessary for a career in psychology? The truth is quite different. Professional psychologists rely on research based on quantitative methods and statistics in their work.
- Furthermore, psychologists in all fields are often expected to do relevant psychological research as part of their work role.
- Many of the professions outside psychology entered by students use knowledge based on quantitative methods and statistics. So a good working knowledge of statistics puts psychology students at an advantage in the employment market.
- Learning statistics can be made hard because psychologists often employ old and outmoded statistical ideas. Some of these ideas are not only unhelpful but also unworkable. This helps contribute to the fog of confusion surrounding statistics. Textbook writers are frequently guilty of perpetuating these counterproductive ideas.
- Too much emphasis is placed on significance testing. Worthwhile as this is, statistics can contribute much more to research than just that. It is important to have an overview of the extensive contribution that statistics makes to psychological knowledge.
- Not many mathematical skills are needed to develop a good working understanding of the role of statistics in psychological research. All but a few students have these skills. Even where these skills have got a little rusty, they can be quickly relearnt by motivated students.

1.1 Introduction

For many psychology students the formula is simple: statistics = punishment. Statistics is 'sadistics'. Most would avoid statistics given the choice. This makes a very unpromising learning environment. And what about the poor soul teaching statistics to reluctant students? Student ratings of statistics modules can bring tears to the eyes of all but the most classroom weary and hardened of professors and lecturers. All round, what could be more unsatisfactory? Couldn't statistics simply be left out of psychology degrees? Well yes, but it is unlikely to happen. Statistics is central to quantitative research in psychology and the creation of psychological knowledge. Surely there are many practitioners who do a great deal of good without needing statistics? Even if this were true in the past it is not so nowadays. The rigid distinction between researcher and practitioner no longer applies. Modern practitioners combine practice with research. Psychologists working in the prison service, in clinical psychology, in education and so forth are usually expected to do some research. This is also true for many of the other professions that psychology graduates may enter. We are living in an information-based society and a great deal of this comes from statistically based research. The bottom line is that some knowledge of statistics is professionally important.

However, statistics (along with mathematics) is generally negatively evaluated. The average person has an attitude to statistics without knowing much about what it involves. They may 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. Statistics can be used misleadingly but that is not generally the objective. We all know that minor adjustments to a graph can distort the truth. A modest growth or decline in a graph may be dramatically changed to seem miraculous or calamitous. Statistics deserves greater respect than its reputation suggests.

The word statistics comes from the Latin for state (as in nation). Statistics originally was the information collected by the State to help Government in its decision-making. The Government's appetite for statistical information is prodigious as we all know. All areas of the Government's planning and decision making are guided by statistical data – pay, pensions, taxes, health services, prisons, the police and so forth. Big supermarkets use it, charities use it, the health service uses it, industrialists use it – you name it and they probably use statistics.

Sound statistical knowledge is fundamental to understanding, planning and analysing research. Nevertheless, students study psychology to know about psychology – not to study statistics. They may not realise that the psychology that they will learn is very dependent upon statistics. Of course there is qualitative research in psychology which does not involve statistics almost by definition but qualitative research is very much in the minority. For the foreseeable future, quantitative methods are likely to have a strong grip on the bulk of psychological research. Statistics and psychology are intertwined.

Statistics isn't taught just to punish students – no matter if it feels that way. It is central to the whole enterprise of psychology. So why not try to see statistics as 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.

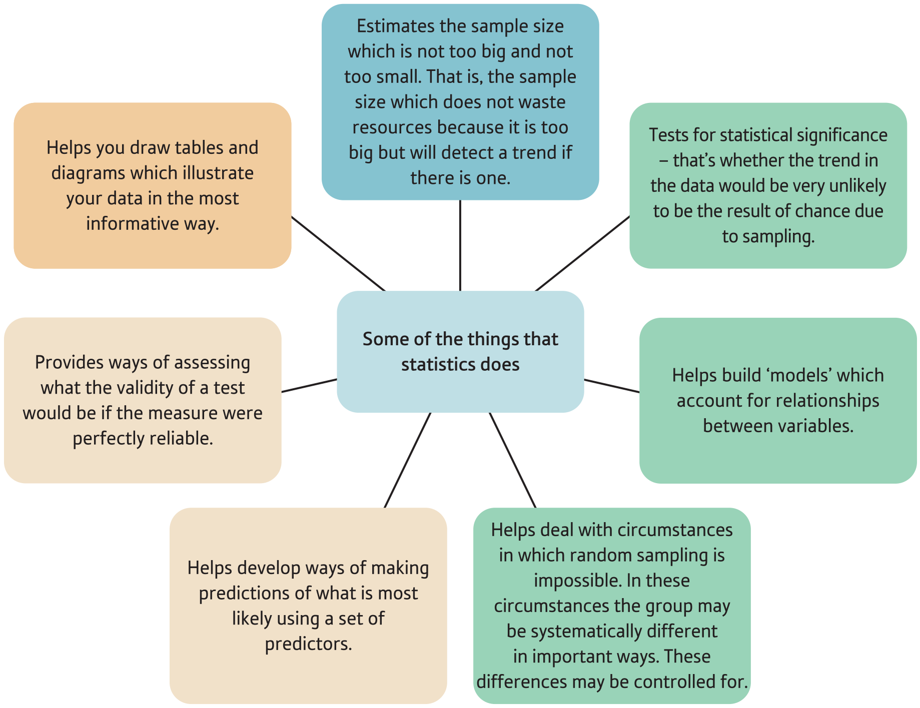


FIGURE 1.1 Some things statistics can do for the researcher

1.2 Research on learning statistics

Our culturally endemic negative view of statistics ensures that the research on psychology students and statistics is generally depressing reading. Trepidation and anxiety just about sum up the initial response of students to learning statistics. Gordon (2004) surveyed a large number of Australian psychology students about their experience of statistics classes. Three-quarters would not have studied it except it was compulsory. They saw statistics as boring and difficult and felt that psychology and psychologists do not need statistics. Their approach was to treat statistics like it was a few mechanical procedures to be applied without understanding 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. (paragraph numbered 18)

Those students who tried to master the methods and concepts of statistics nevertheless had difficulty in understanding its importance. Students 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 negative group were not generally less able students and did not generally do worse than other students on other modules. 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.

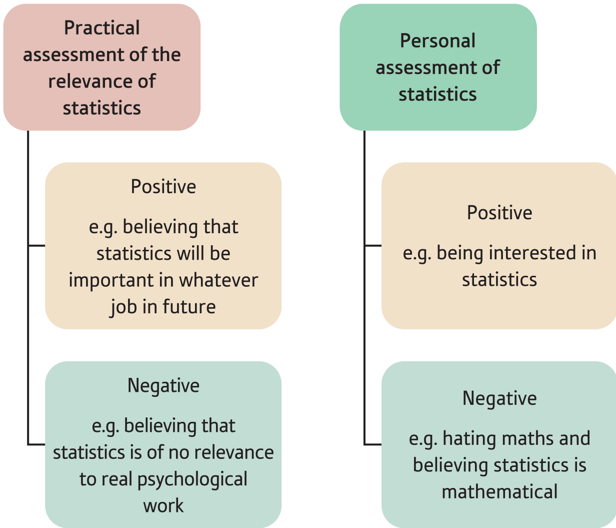


FIGURE 1.2 Responses of students to statistics according to Gordon (1995)

1.3 What makes learning statistics difficult?

It is usually recognised by university staff that teaching statistics involves dealing with the anxieties, beliefs and negative attitudes concerning the subject (Schau, 2003). Background issues like these may be the most important things in the learning process. University can be an experience full of emotion, and emotion affects learning. This is perhaps most 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 but real. Being at university and studying statistics follows a long period of personal development through schooling (and for some at work). Personal histories, experiences, needs and goals are reflected in our strategies for coping with statistics (Gordon, 2004). These influence the way that we think about our learning processes and education more generally. Beliefs such as 'I'm no good at maths' will impact on our response to statistics.

In other words, a student can bring to learning statistics baggage which may seriously interfere with studying. Issues to do with one's mathematical ability are high on the list. Some students may (incorrectly) assume that their low maths skills make statistics too hard for them. This is reinforced by those departments which require good maths grades for admission. With other time pressures, such students may adopt avoidance tactics such as skipping lectures rather than putting the time into studying statistics. Furthermore, every statistics class has its own culture in which students influence each other's attitudes to learning statistics. A class dominated by students antagonistic to statistics is not a good learning environment. Acting silly, talking in class or plagiarising the work of other students just does not help.

Whether mathematical ability is important to making a good statistics student is doubtful. 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. Anxiety plays its part primarily through a specific form of anxiety known as mathematics (math) anxiety.

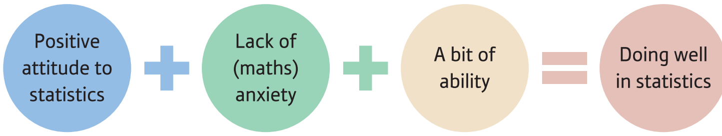


FIGURE 1.3 Formula for doing well in statistics based on research findings

This is more important in this context than trait 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 worse 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. That is, in itself, poor mathematical ability is not primarily a cause of worse results.

If more research evidence is needed, using a formal measure known as the Survey of Attitudes toward Statistics, Zimprich (2012) showed that 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 were interrelated, as one might expect. They also correlated with actual achievement in statistics. These attitudes were much more important than actual maths ability in terms of how well students do 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.

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. Generally speaking, there are few occasions when it is necessary to do calculations by hand and then these are usually simple. Often you will find websites which will calculate the things which SPSS does not do. Mostly, though, the statistical analyses you need are available on SPSS and other statistics programs. So long as you have entered your data correctly and chosen an appropriate statistical analysis you do not have to worry about the calculation. Some basic mathematics is helpful, of course, when learning about statistics since numbers and symbols won't be quite so daunting. Statistics is a maths-based discipline and its concepts are generally defined by formulae rather than in words. So if you are good at understanding mathematical formulae then this is an advantage, though far from necessary. Even researchers differ widely in their mathematical skills and many would not see themselves as mathematical at all. Yet they have learnt to use statistics appropriately and intelligently, which is very much the task facing students. You need to understand the purpose of a statistical test and why it was developed, understand a little about how it works, know when to use it and most of all be able to make sense of the computer output. Maths is peripheral for the most part.

Just what mathematical knowledge does one need to achieve a working knowledge of statistical analysis? By and large if you understand the concepts of addition, subtraction, multiplication and division then you have the basics. You may not always get the right answer but the important thing is that you understand what these mathematical operations are about. What might you need beyond this? Probably just the following:

- 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 short time spent 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. All in all, the requirements are not very demanding. Anything that has been forgotten or never learnt will be quickly picked up by a motivated student. Not all lecturers will share this opinion. Nevertheless, the overwhelming majority know that students can really struggle with statistics for any number of reasons. So they provide teaching which serves the needs of all students taking the psychology programme, not just the maths-able ones.

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 and Garfield, 2004; Ruggeri, Dempster and Hanna, 2011). It has its own language and concepts. Grasping the statistical way of thinking and learning to speak statistical language takes some effort. Students in all sorts of disciplines struggle somewhat with statistics, it is not just psychology students. Statistical thinking is a different way of thinking.

Broadly speaking, different research designs require different statistical techniques. So you really need to understand the different kinds of research design before statistical analysis makes sense. Statistical problems in research are often research design problems. You really do have to formulate your research question, your hypotheses and your research design carefully for the statistical analysis to fall into place. Every degree course will give you a grounding in research methods and how research is done. But such knowledge will not translate directly into an ability to do research. This is developed through practical or lab classes in which you experience the process of doing research. Although research skills build up quite slowly over the course of your degree these skills are little or nothing to do with mathematics. They are about the application of logic and thought to the research process. Statistical analysis takes a minor role compared to the more general research skills involved in a quantitative study. If you are confused about your research question, your hypotheses and your research design, it follows that you will be confused about the appropriate statistical analysis.

1.4 Positive about statistics

So how does one go about having a more positive attitude towards statistics? Part of 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. Just why did statistics become so important in modern research when for centuries people did experiments and other research without significance testing and the like? One of the most well-known statistical techniques used by psychologists is the *t*-test (see Chapters 13 and 14) or the Student *t*-test as it is also known. For decades, psychology students have learnt to do a *t*-test. 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.

One issue that was important to the company was 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. Gosset worked on the problem of the extent of error that is likely to occur when small samples were being used in quality control. He developed a mathematical way of calculating the likely error which can occur when testing samples compared to the entire output. If you decided to take a sample of just 10 bottles, to what extent is the sample likely to mislead you about the alcoholic strength of the product in general?

Of course, you will never know from a sample exactly what the error will be but Gosset was able to estimate its likely extent from the variability within the sample. Put into a formula, this is the idea of standard error which plagues many students on introductory statistics courses. The *t*-test is based on standard error. 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 many 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. A test of statistical significance addresses the possibility that a trend that we find in our sample could simply have occurred by chance when there is no trend in reality. 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.) But significance testing is only a small part of statistics, which provides a whole range of tools to help researchers (and students) address the practical problems of data analysis. Research data can be very simple but also very complex. Statistics helps sort out the complexity and uncertainty involved in understanding your data.

What sample size do I need?

Gosset's focus on small samples begs the question of how small a sample can be used. 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 inadvertently all of the time simply because researchers (including students) do not address the question of sample size properly. Often the advice given to those asking what sample size to use is that they should get as big a sample as they can. But this is a crude way of going about deciding sample size. Even the smallest trend will be statistically significant if the sample size is large enough. However, there is little point in using large samples when smaller ones would be adequate. The optimum sample size depends on the size of the effect the researcher thinks is worthwhile investigating, the statistical significance level required and the risk of not supporting the hypothesis when it is in fact true that the researcher is prepared to take. There are conventional values for the latter two but the researcher may wish to vary these.

There are no objective criteria which tell us what potential size of effect is worth studying which apply irrespective of circumstances. It might appear obvious that research should prioritise large trends but it is not as simple as that. In medical research, for instance, there are examples of very small trends which nevertheless save lives. Taking aspirin has a small effect on reducing the risk of heart attacks but saves lives in aspirin takers compared with a control group. The size of a trend worth the research effort therefore depends on what is being considered. A pill which prevents cancer in 10% of people would be of more interest than a pill which prevents flatulence in 10% of people, for example. So if a researcher designs a study which has a sample size too low to establish a statistically significant trend then this would be more worrisome in the case of the cancer cure than in the case of the flatulence cure. Chapter 40 explains how to go about deciding sample size in a considered, rational way. This area of statistics is known as *statistical power analysis*. So the apparently simple question of the sample size needed is rather more complex than at first appears.

This is not the place to give a full overview of the role of statistics in psychological research. It is important, though, to stress that statistics can help you with your research in many ways. This is hardly surprising since statisticians seek to address many of the problems which researchers face in their quantitative research. Now this book is just about as comprehensive as understandable statistics texts get but not everything that statistics can do is represented. Nevertheless, you will find a great deal which goes far beyond the issue of statistical significance. Take, for example, factor analysis (Chapter 33). This is not at all about statistical significance but a way of finding or identifying the basic dimensions in your data. So, for example, many famous theories of personality and theories of intelligence have emerged out of factor analysis – for instance, that of Hans Eysenck (Eysenck and Eysenck, 1976) which suggests that extraversion, neuroticism and psychoticism are the major underlying dimensions or components of personality on which people differ. There is no way that a researcher can simply look at their data, which can be enormously complex, and decide what its underlying structure is. It is not possible to identify extraversion, neuroticism and psychoticism simply by looking at the data from a 50-item questionnaire that has been completed by 2000 participants. But statisticians (and psychologists with a strong interest in statistics) developed methods of doing just that and computers make this as simple as it can be.

Statistics also has a very important role in model building. This sounds complicated but it isn't too difficult. A model is simply a proposed set of relationships between variables. So, for instance, the relationships shown in Figure 1.3 between various characteristics of students studying statistics and their achievement in tests and examinations is a sort of model. Statistics addresses just how well the data fits the proposed model – there may be other characteristics of the student that need to be considered in addition to those in Figure 1.3 in order to account fully for how well students do in statistics. The researcher may propose models but, equally, statistical techniques also help identify potential models.

Some of the other things which statistics can help you with include:

- Is the trend that I have just found in my data big or small?
- Does this line of research show potential for further development?
- Are the measures that I am using sufficiently reliable and valid to detect a trend that I am interested in?

- Is it possible to amalgamate a number of variables into a single, more readily understood one?
- Can I eliminate competing explanations of my findings so as to give more credence to my hypothesis?
- How best can I present my data graphically in order to visually present my findings to an audience at a conference?
- Can I combine the findings of different studies so as to have a good idea of the typical findings of past research?

Statistics is just one aspect of the decision-making process which underlies research in psychology. It should not dominate a researcher's thinking exclusively. It is not even the most important part of research. But without it your decision-making may not be optimal.

1.5 What statistics doesn't do

Years of experience teaching statistics means, of course, that we were the statistics doctors whom students having problems with analysing their data came to – or even got sent to. These encounters vary widely. Some students simply do not have a clue about statistics and cannot relate what they learnt in statistics lectures with their own research. Other students appear to want help but really they are seeking confirmation that their ideas for their analysis are correct or that they have understood their data correctly. Yet others have designed their research so badly that either it is difficult to analyse at all or it is difficult to analyse using the statistics that the student knows at this point.

You should not blame your lack of statistical knowledge when your research does not allow you to answer the question that you set about addressing in the research plan. It is essential to think carefully about what your research design achieves prior to collecting data. While planning your research, ask yourself just how you will answer your research questions using the data you are collecting. The less clear you are about your research questions then the more difficult this is to do. And your lack of statistical knowledge will rarely be the problem.

It is surprising the number of students who stumble early on in the research process like this. Deadlines for research proposal submissions can result in the writing of a research plan which is not as good or clear as it should be. You should be in a position to plan your analysis in advance of collecting your data. Just how will you go about doing your analysis? This implies that you could insert more or less random numbers, etc. into your analysis and go on to perform the analysis based on these before you collect your actual data. What tables would you need? What statistical techniques would be employed? Such questions ought to be thought about very early in the planning of one's research. But the temptation is to leave the statistical issues to last in the hope that something or someone will come to your rescue. Such pre-planning is a hard thing to do as a beginner but if you cannot detail your analysis early on then why do you expect to be hit by a wave of insight after you have collected your data?

So sometimes students do not have a clear grasp of the research that they are proposing to do. Confusion can be caused by trying to achieve too much in one study, but insufficient preparation may also be responsible. It is difficult for any of us to be clear about our ideas without investing the time to think carefully. You should talk to anyone prepared to listen. There is no quicker way of recognising problems with your research proposals than finding yourself unable to explain clearly to someone else just what you