

Preface



My aim was to produce a statistics book with two characteristics: to assume that the reader is using a computer to analyse data and to contain absolutely no equations.

This is a handbook for biologists who want to process their data through a statistical package on the computer, to select the most appropriate methods and extract the important information from the, often confusing, output that is produced. It is aimed, primarily, at undergraduates and masters students in the biological sciences who have to use statistics in practical classes and projects. Such users of statistics don't have to understand exactly how the test works or how to do the actual calculations. These things are not covered in this book as there are more than enough books providing such information already. What is important is that the right statistical test is used and the right inferences made from the output of the test. An extensive key to statistical tests is included for the former and the bulk of the book is made up of descriptions of how to carry out the tests to address the latter.

In several years of teaching statistics to biology students it is clear to me that most students don't really care how or why the test works. They do care a great deal that they are using an appropriate test and interpreting the results properly. I think that this is a fair aim to have for occasional users of statistics. Of course, anyone going on to use statistics frequently should become familiar with the way that calculations manipulate the data to produce the output as this will give a better understanding of the test.

If this book has a message it is this: think about the statistics *before* you collect the data! So many times I have seen rather distraught students unable to analyse their precious data because the experimental design they used was inappropriate. On such occasions I try to find a compromise test that will make the best of a bad job but this often leads to a weaker conclusion than might have been possible if more forethought had been applied from the outset. There is no doubt that if experiments or sampling strategies are designed with the statistics in mind better science will result.

Statistics are often seen by students as the 'thing you must do to data at the end'. Please try to avoid falling into this trap yourself. Thought experiments producing dummy data are a good way to try out experimental designs and are much less labour-intensive than real ones!

Although there are almost no equations in this book I'm afraid there was no way to totally avoid statistical jargon. To ease the pain somewhat, an extensive Glossary and key to symbols are included. So when you are navigating your way through the key to choosing a test you should look up any words you don't understand.

In this book I have given extensive instructions for the use of four commonly encountered software packages: SPSS, R, Excel and MINITAB. However, the key to choosing a statistical test is not at all package-specific, so if you use a software package other than the four I focus on or if you are using a calculator you will still be able to get a good deal out of this book.

If every sample gave the same result there would be no need for statistics. However, all aspects of biology are filled with variation. It is statistics that can be used to penetrate the haze of experimental error and the inherent variability of the natural world to reach the underlying causes and processes at work. So, try not to hate statistics, they are merely a tool that, when used wisely and properly, can make the life of a biologist much simpler and give conclusions a sound basis.

The third edition

In the 8 years since I wrote the second edition of this book there have, of course, been several new versions of the software produced. I have received many comments about the previous editions and I am grateful for the many suggestions on how to improve the text and coverage. Requests to add further statistical packages have been the most common suggestion for change. There was surprisingly little consensus on the packages to add for the second edition, but since 2000 the freely available, and very powerful, package R has become extremely widely used so I have added that to the mix this time.

How to use this book

This is definitely not a book that should be read from cover to cover. It is a book to refer to when you need assistance with statistical analysis, either when choosing an appropriate test or when carrying it out. The basics of statistical analysis and experimental design are covered briefly but those sections are intended mostly as a revision aid, or to outline of some of the more important concepts. The reviews of other statistics books may help you choose those that are most appropriate for you if you want or need more details.

The heart of the book is the key. The rest of the book hinges on the key, explaining how to carry out the tests, giving assistance with the statistical terms in the Glossary or giving tips on the use of computers and packages.

Packages used

MINITAB® version 15, MINITAB Inc.

SPSS® versions 16 and 17, SPSS Inc.

Excel™ version 2007 and 2008 for Mac, Microsoft Corporation

Running on:

Windows® versions XP, 2000, 7 and Vista, Microsoft Corporation

Mac OS 10, Apple Inc.

Example data

In the spirit of dummy data collection, all example data used throughout this book have been fabricated. Any similarity to data alive or dead is purely coincidental.

Acknowledgements for the first edition

Thanks to Sheena McNamee for support during the writing process, to Andrea Gillmeister and two anonymous reviewers for commenting on an early version of the manuscript and to Terry Crawford, Jo Dunn, David Murrell and Josephine Pithon for recommending and lending various books. Thanks also to Ian Sherman and Susan Sternberg at Blackwell and to many of my colleagues who told me that the general idea of a book like this was a sound one. Finally, I would especially like to thank the students at the University of York, UK, who brought me the problems that provided the inspiration for this book.

Acknowledgements for the second edition

Thanks to all the many people who contacted me with suggestions and comments about the first edition. I hope you can see that many of the corrections and improvements have come directly from you. Five anonymous reviewers provided many useful comments about the proposal for a second edition. Thanks to Sarah Shannon, Cee Brandston, Katrina McCallum and many others at Blackwell for seeing this book through and especially for producing a second superb and striking cover. S'Albufera Natural Parc and Nick Riddiford provided a very convenient bolt-hole for writing. Once again, I give special thanks to Sheena and to my colleagues, PhD students and undergraduate students at the University of York. Finally, thanks to everyone on the MRes EEM course over the last 4 years.

Acknowledgements for the third edition

It's been thanks to the pushing of Ward Cooper at Wiley-Blackwell and Sheena McNamee that this third edition has seen the light of day. Thanks to Emma Rand, Olivier Missa and Frank Schurr for encouraging me to enter the brave new world of R. Thanks to Nik Prowse for guiding me through the final editing.

Calvin Dytham,
York 1998, 2002 and 2010