palgrave macmillan

Serious Stats

A Guide to Advanced Statistics for the Behavioral Sciences

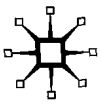
Thom Baguley

SERIOUS STATS

A GUIDE TO ADVANCED STATISTICS FOR THE BEHAVIORAL SCIENCES

THOM BAGULEY
Professor of Experimental Psychology, Nottingham Trent University, UK

palgrave
macmillan



First published 2012 by
PALGRAVE MACMILLAN

Palgrave Macmillan in the UK is an imprint of Macmillan Publishers Limited, registered in England, company number 785998, of Houndmills, Basingstoke, Hampshire RG21 6XS.

Palgrave Macmillan in the US is a division of St Martin's Press LLC, 175 Fifth Avenue, New York, NY 10010.

Palgrave Macmillan is the global academic imprint of the above companies and has companies and representatives throughout the world.

Palgrave® and Macmillan® are registered trademarks in the United States, the United Kingdom, Europe and other countries.

ISBN 978–0–230–57717–6 hardback
ISBN 978–0–230–57718–3 paperback

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources. Logging, pulping and manufacturing processes are expected to conform to the environmental regulations of the country of origin.

A catalogue record for this book is available from the British Library.

A catalog record for this book is available from the Library of Congress.

10 9 8 7 6 5 4 3 2 1
21 20 19 18 17 16 15 14 13 12

Printed in China

SERIOUS STATS

Preface

About this book

This book is a bridging text for students and researchers in the human and behavioral sciences. The ideal reader will have some familiarity with inferential statistics – perhaps as part of an undergraduate degree in a discipline such as psychology, cognitive science or ergonomics – and be interested in deepening their understanding or learning new material. This book aims to bridge the gap between a reader's existing understanding of statistics and that required to apply and interpret more advanced statistical procedures.

I have also tried to make the book a helpful resource for experienced researchers who wish to refresh their statistical knowledge or who have good understanding of a 'narrow' but fairly advanced topic such as analysis of variance. I hope it will allow these readers to expand from islands of existing expertise to new territory.

The book starts with a review of basic inferential statistics, beginning with descriptive statistics, probability distributions and statistical inference (in the form of confidence intervals and significance tests). If you are already familiar with these topics I would encourage you to look through these chapters to refresh your understanding. In addition, this material may be presented in a slightly different way (e.g., from a different perspective or in greater depth).

Later chapters introduce core topics such as, regression, correlation and covariance, effect size, and statistical power. Unless you have advanced training in statistics it is likely that you will benefit from looking closely at this material – it is fundamental to an appreciation of later content. Two further chapters consider the messiness inherent in working with real data (particularly data from human participants). The approach I adopt is to give a taster of some methods for exploring and dealing with messy data, rather than provide a comprehensive recipe for checking and solving every possible problem. This is both for practical reasons (as each of these chapters could be a book in its own right) and because the best approach in any particular situation depends on what you are trying to do and the context from which the data are drawn.

Later chapters cover what I consider to be advanced material: multiple regression, analysis of variance, analysis of covariance, and the general linear model. Before covering these topics I review alternatives to classical, frequentist inference (and significance tests in particular). In order to get the most out of the more advanced material in the book, you will need to understand the problems inherent in relying (solely) on a p value from a significance test for inference. I also think it important to go beyond criticism of the p value approach and present viable alternatives. Three are presented here: Bayesian, likelihood, and information theoretic approaches to inference. There are important connections (and distinctions) between these three approaches. In this chapter, I sacrifice depth for breadth (though there is sufficient material to run a range of analyses using each approach).

The final chapters explore the most challenging topics. Also included are chapters on interaction effects and contrasts. These topics are extremely important for researchers in the human and behavioral sciences, but are often covered only briefly (if at all) in introductory classes. My goal here is to remedy this deficit. The final two chapters introduce generalized linear models

(for discrete outcomes) and multilevel models (with emphasis on repeated measures models). I have tried to emphasize the links between these advanced topics and the general linear model and to demonstrate what they offer over and above simpler models.

If you want to learn and understand what is covered, it is essential that you have a go at applying it. Each chapter contains worked examples. Many of these use real data sets. These are necessarily a bit messy and don't always lead to clear answers (and several data sets are chosen because they have interesting or unusual quirks). My aim is to illustrate some of the challenges of working with real data sets (and the importance of data exploration and model checking). In other cases I have resorted to creating artificial data sets to illustrate a particular point, or to make it easier to conduct calculations by hand. These data sets are carefully constructed to meet the requirements of the example – though you will sometimes encounter real data sets with similar properties. In general, the early examples use hand calculation while later examples require you to use a computer.

Hand calculation can sometimes help you to understand how an equation works or demystify a (supposedly) complex technique. This will depend on your confidence and ability with basic mathematical operations. This doesn't work with every procedure, and in many examples I explain how to use a computer package to provide intermediate values that, when put together in the right way, illustrate what is going on. From time to time the mathematics is sufficiently challenging that I merely describe the gist of what is happening (and rely on the computer to provide a complete solution). Where necessary, I refer interested readers to a more detailed mathematical account of what is taking place.

The contents of the book differ from the coverage of a typical introductory or intermediate statistics course in the behavioral or human sciences. One difference is the breadth of coverage, which runs from descriptive statistics to generalized linear and multilevel models. Another is the reduced emphasis on null hypothesis significance tests and increased emphasis on confidence intervals or other inferential tools. Several topics have more prominence than you might expect: graphical methods; effect size; contrasts and interactions. Other topics have less emphasis or are presented differently: psychometrics, multivariate analysis of variance, non-parametric statistics; and pairwise comparisons. I have chosen to focus on univariate methods – methods for the analysis of a single outcome measure (though there may be many predictor variables). Covering multivariate statistics (in the sense of modeling multiple outcomes) and psychometrics would probably have doubled the page count. Nonparametric statistics are covered, but in an atypical way. Several methods, often considered to be 'nonparametric' (e.g., bootstrapping, kernel density estimation, the rank transformation and robust regression), are integrated into the text at appropriate points. I have, in particular, avoided describing a large number of rank transformation tests in detail. My preference is to emphasize the link between parametric and so-called non-parametric approaches and to encourage consideration of robust methods as alternatives to the usual (e.g., least squares) models.

If there is a single message to take away from this book, it is that statistical modeling is not a set of recipes or instructions. It is the search for a model or set of models that capture the regularities and uncertainties in data, and help us to understand what is going on.

Software

Many of the statistical tools described in the book require specialist software to run them. Nearly all the examples were implemented in the free, open source statistical programming

environment R (R Core Development Team, 2011). This will run on PC, Mac and Linux operating systems. Installation on a Mac or PC is generally very easy. For details on downloading and installing R see: http://cran.r-project.org/

R works slightly differently on PC, Mac and Linux machines and for this reason I have, for the most part, avoided referring to platform-specific features of R (e.g., resizing windows, printing, opening or saving files). There are many online guides and books that run through the basics of installing and running R (in addition to the information available when you download R for the first time).

At the end of each chapter, I provide a detailed description of the R code used to reproduce the examples in that chapter. Data sets and R scripts for each chapter are available with the online resources for this book. I assume no previous knowledge of R (and only limited familiarity with statistical computing). Although I am not trying to teach R *per se*, I have tried to include enough explanation of how R works for readers, if they wish, to learn it as they go along. To this end, the complexity of the R code being used increases gradually from Chapter 1 through to Chapter 18. In some cases I have glossed over fine grain technical details about how R works (e.g., the difference between 'modes' of vector) and used generic terms such as 'string' alongside R-specific terms such as 'data frame'.

As well as R code, I provide very brief notes on relevant SPSS syntax at the end of each chapter (where relevant). SPSS is the most widely used statistics package in the human and behavioral sciences (in the UK at least). These notes are included for two reasons: (1) to reveal some of the hidden features and capabilities of SPSS, and (2) to highlight the advantages of using R alongside or in place of SPSS.

Mathematics

To get the most from this book you will need to have basic mathematical competence. I assume readers will have mastered basic arithmetic (e.g., addition, subtraction, division and multiplication) and be familiar with concepts such as, fractions, decimals, rounding, negative numbers, squares and square roots, and cubes and cubed roots. Knowing the order in which to apply arithmetic operations (e.g., PEMDAS or BODMAS) is also necessary. You should also have some understanding of percentages, probabilities and ratios (and perhaps reciprocals, exponents, logarithms and factorials) and simple algebra. If you are rusty on any of these topics don't worry too much – as there will be some 'refresher' material in each chapter.

I would expect readers to be able to answer the following arithmetic problems without much difficulty:

$$4+3\times 5=? \quad 6(3-1)=? \quad \sqrt{9}=?$$
$$10^2=? \quad \frac{9}{2}=? \quad 0.5\times 0.5=?$$

I would also expect readers to understand what the following equations mean, and (perhaps with a little help) be able to solve them:

$$3x+1=10\Rightarrow x=? \quad 4!=? \quad \sqrt{2}\times\sqrt{2}=?$$
$$P(A)+P(\sim A)=? \quad \sqrt[3]{27}=? \quad -2\times -3=?$$

In addition to mathematical competence, I anticipate some familiarity with data collection, exploration and analysis. This may include experimental design, descriptive statistics and simple graphical methods such as line graphs or scatter plots (and technical terms such as *x*-axis and *y*-axis). Again, the text contains refresher material on most of these topics.

Many of the examples refer to 'hand calculation'. This is a fuzzy concept, but I take it to mean reproducing the calculations step-by-step as if doing them by hand. It therefore includes using paper and pencil, mental arithmetic, pocket calculators or spreadsheets (if used in the correct way). In fact, a spreadsheet is one of the best ways to organize hand calculations to understand what they are doing (provided you know how to use one and are careful to set out all the intermediate steps).

Boxed sections and online supplements

As well as learning features (such as examples, sections on R code, and SPSS notes), there are two types of boxed section used throughout the book. One type covers key concepts or important ideas that are referred to in several different chapters. These are referred to as 'key concepts' and numbered by chapter and serial position within chapter (e.g., Key Concept 2.1 is the first key concept box in Chapter 2). The other is a more traditional boxed section that is used to improve the flow of the text and contains material that is relatively self-contained (and that is generally referred to again only within that chapter). These are referred to as 'boxes' and numbered by chapter and serial position within chapter (e.g., Box 1.2 is the second boxed section in Chapter 1).

In addition to the boxed sections there are five online supplements. Supplements 1, 2 and 5 cover advanced topics that are not central to the text, but will be very useful for some readers (meta-analysis, dealing with missing data and loglinear models). Supplements 3 and 4 provide more detail on peripheral topics that are mentioned in the main text (replication probabilities and pseudo-R^2 measures).

Acknowledgments

I would like to thank the following people who have contributed suggestions for the book, read chapters or provided encouragement: Jaime Marshall; Andrew Dunn; Danny Kaye; Mark Andrews; Stanislav Kolenikov; Dave Atkins; Tim Wells; Tom Dunn; Mark Torrance; James Houston and Zoltan Dienes.

THOM BAGULEY

Contents overview

Contents

List of tables

List of figures