Essential Statistics for the Pharmaceutical Sciences

Philip Rowe

Liverpool John Moores University, UK

BICENTENNIAL 1807 WILEY 2007 BICENTENNIAL

Essential Statistics for the Pharmaceutical Sciences

Essential Statistics for the Pharmaceutical Sciences

Philip Rowe

Liverpool John Moores University, UK

Telephone (+44) 1243 779777

Email (for orders and customer service enquiries): cs-books@wiley.co.uk
Visit our Home Page on www.wileyeurope.com or www.wiley.com

Other Wiley Editorial Offices

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA
Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA
Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany
John Wiley & Sons Australia Ltd, 33 Park Road, Milton, Queensland 4064, Australia
John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01, Jin Xing Distripark, Singapore 129809
John Wiley & Sons Canada Ltd, 6045 Freemont Blvd, Mississauga, Ontario, L5R 4J3

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Anniversary Logo Design: Richard J. Pacifico

Library of Congress Cataloguing-in-Publication Data

Rowe, Philip.
Essential statistics for the pharmaceutical sciences / Philip Rowe.
p. ; cm.
Includes bibliographical references and index.
ISBN-13: 978-0-470-03470-5 (cloth : alk. paper)
ISBN-13: 978-0-470-03468-2 (pbk. : alk. paper)
1. Drugs–Research–Statistical methods. 2. Pharmacology–Statistical methods. I. Title.
[DNLM: 1. Pharmacology–methods. 2. Statistics. QV 20.5 R879e 2007]
RS57.R69 2007
615'.1072–dc22 2006102028

British Library Cataloguing in Publication Data

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

ISBN-13: 978 0-470-03470-5 (HB)
ISBN-13: 978 0-470-03468-2 (PB)

Typeset in 10.5/12.5pt Minion by Thomson Digital
Printed in Great Britain by Antony Rowe Ltd., Chippenham, Wiltshire
This book is printed on acid-free paper responsibly manufactured from sustainable forestry in which at least two trees are planted for each one used for paper production.

To

Carol, Joshua and Nathan

Contents

Preface

At whom is this book aimed?

Statistics users rather than statisticians

As a subject, statistics is boring, irrelevant and incomprehensible. Well, traditionally it is anyway. The subject does not have to be that bad; It simply has not escaped from a self-imposed time-warp. Thirty years ago, there were no easy-to-use computer packages that would do all the hard sums for us. So, inevitably, the subject was heavily bogged down in the 'How to' questions – how do we most efficiently calculate the average weight of 500 potatoes? How do we work out whether the yield of turnips really did increase when we used a couple of extra shovels of horse manure? How should we calculate whether there really is a relationship between rainfall and the size of our apples? Thirty years ago, statistical authors had a genuine excuse for their books being clogged up with detailed methods of calculation. However, one might ask, what is their excuse today? Nowadays, nobody in their right mind would manually calculate a complex statistical test, so why do they still insist on telling us how to do it?

Clearly the world does need genuine 'statisticians' to maintain and improve analytical methods and such people do need to understand the detailed working of specific procedures. However, the majority of us are not statisticians as such; we just want to use statistical methods to achieve a particular end. We should be distinguishing between 'statisticians' and 'statistics users'. It is rather like the distinction between automotive engineers and motorists. Automotive engineering is an honourable and necessary profession, but most of us are just simple motorists. Fortunately, within the field of motoring, the literature does respect the difference. If a book is written for motorists, it will simply tell us that, if we want to go faster, we should press the accelerator. It will not bore us stiff trying to explain how the fuel injection system achieves that particular end. Unfortunately that logic has never crept into statistics books. The wretched things still insist on trying to explain the internal workings of the Kolmogorov–Smirnov test to an audience who could not care less.

Well, good news! This book is quite happy to treat any statistical calculation as a black box. It will explain what needs to go into the box and it will explain what comes out the other end, but you can remain as ignorant about what goes on inside the box as you are about how your power steering works. This approach is not just lazy or negative. By stripping away all the irrelevant bits, we can focus on the aspects that actually matter. This book will try to concentrate on those issues that statistics users really do need to understand:

- Why are statistical procedures necessary at all?
- How can statistics help in planning experiments?
- Which procedure should I employ to analyse the results?
- What do the statistical results actually mean when I have got them?

Who are these 'statistics users'?

The people that this book is aimed at are those thousands of people who have to use statistical procedures without having any ambition to become statisticians. There are any number of student programmes, ranging from pharmacology and botany through to business studies and psychology, which will include an element of statistics. These students will have to learn to use at least the more basic statistical methods. There are also those of us engaged in research in academia or industry. Some of us will have to carry out our own statistical analyses and others will be able to call on the services of professional statisticians. However, even where professionals are to hand, there is still the problem of communication. If you do not even know what the words mean, you are going to have great difficulty explaining to a statistician exactly what you want to do. The intention is that all of the above should find this book useful.

If you are a statistics student or a professional statistician, my advice would be to put this book down now! You will probably find its dismissive attitude towards the mathematical basis of your trade extremely irritating. Indeed, if at least one traditional statistician does not complain bitterly about this book, I shall be rather disappointed.

To what subject area is the book relevant?

All the statistical procedures and tests mentioned are illustrated with practical examples and data sets. The cases are drawn from the pharmaceutical sciences and this is reflected in the book's title. However, pretty well all the methods described and the principles explored are perfectly relevant to a wide range of scientific research, including pharmaceutical, biological, biomedical and chemical sciences.

At what level is it aimed?

The book is aimed at undergraduate science students and their teachers and less experienced researchers.

The early chapters (1–5) are fairly basic. They cover data description (mean, median, mode, standard deviation and quartile values) and introduce the problem of describing uncertainty due to sampling error (SEM and 95 per cent confidence interval for the mean). In theory, much of this should be familiar from secondary education, but in the author's experience, the reality is that many new students cannot (for example) calculate the median for a small data set. These chapters are therefore relevant to level 1 students, for either teaching or revision purposes.

Chapters 6–17 then cover the most commonly used statistical tests. Most undergraduate programmes will introduce such material fairly early on (levels 1 or 2). The approach used is not the traditional one of giving equal weight to a wide range of techniques. As the focus of the book is the various issues surrounding statistical testing rather than methods of calculation, one test (the two-sample *t*-test) has been used to illustrate all the relevant issues (Chapters 6–11). Further chapters (12–17) then deal with other tests more briefly, referring back to principles that have now been established.

The final chapters (18 and 19) cover some real-world problems that students probably will not run into until their final year, when carrying out their own independent research projects. The issues considered are multiple testing (all too common in student projects!) and questionnaire design and analysis. While the latter does not introduce many fundamentally new concepts, the use of questionnaires has increased so much, it seemed useful to bring together all the relevant points in a single resource.

Key point and pirate boxes

Key point boxes

Throughout the book you will find key point boxes that look like this:

Proportions of individuals within given ranges

For data that follow a normal distribution:

- about two-thirds of individuals will have values within 1 SD of the mean;
- about 95 per cent of individuals will have values within 2 SD of the mean.

These never provide new information. Their purpose is to summarize and emphasize key points.

Pirate boxes

You will also find pirate boxes that look like this:

> **Switch to a one-sided test after seeing the results**
>
> Even today, this is probably the best and most commonly used statistical fiddle.
>
> You did the experiment and analysed the results by your usual two-sided test. The result fell just short of significance (*P* somewhere between 0.05 and 0.1) There is a simple solution, guaranteed to work every time. Re-run the analysis, but change to a one-sided test, testing for a change in whatever direction you now know the results actually suggest.
>
> Until the main scientific journals get their act into gear, and start insisting that authors register their intentions in advance, there is no way to detect this excellent fiddle. You just need some plausible reason why you 'always intended' to do a one-tailed test in this particular direction, and you are guaranteed to get away with it.

These are written in the style of Machiavelli, but are not actually intended to encourage statistical abuse. The point is to make you alert for misuses that others may try to foist upon you. Forewarned is forearmed. The danger posed is reflected by the number of skull and cross-bone symbols.

Minor hazard. Abuse easy to spot or has limited potential to mislead.

Moderate hazard. The well-informed (e.g. readers of this book) should spot the attempted deception.

Severe hazard. An effective ruse that even the best informed may suspect, but never be able to prove.

Fictitious data

Throughout this book, all the experiments described are entirely fictitious as are their results. Generally the structure of the experiments and the results are realistic. In a few cases, both the structure of an experiment and consequently the analysis of the results may be somewhat simpler than would often be seen in the real world. Clarity seems more important than strict realism.

At several points, judgements are quoted as to how greatly a measured end-point would need to change, for that change to be of any practical consequence. The values quoted are essentially arbitrary, being something that appears realistic to the author. Hopefully these are reasonable estimates, but none of them should be viewed as expert opinion.

A potted summary of this book

The book is aimed at those who have to use statistics, but have no ambition to become statisticians *per se*. It avoids getting bogged down in calculation methods and focuses instead on crucial issues that surround data generation and analysis (sample size estimation, interpretation of statistical results, the hazards of multiple testing, potential abuses, etc.). In this day of statistical packages, it is the latter that cause the real problems, not the number-crunching.

The book's illustrative examples are all taken from the pharmaceutical sciences, so students (and staff) in the areas of pharmacy, pharmacology and pharmaceutical science should feel at home with all the material. However, the issues considered are of concern in most scientific disciplines and should be perfectly clear to anybody from a similar discipline, even if the examples are not immediately familiar. Material is arranged in a developmental manner. Initial chapters are aimed at level 1 students; this material is fairly basic, with special emphasis on random sampling error. The next section then covers key concepts that may be introduced at levels 1 or 2. The final couple of chapters are most likely to be useful during final year research projects; These include one on questionnaire design and analysis.

The book is not tied to any specific statistical package. Instructions should allow readers to enter data into any package and find the key parts of the output. Specific instructions for performing all the procedures, using Minitab or SPSS, are provided in a linked web site (www.staff.ljmu.ac.uk/phaprowe/pharmstats.htm).

Statistical packages

There are any number of statistical packages available. It is not the intention of this book to recommend any particular package. Quite frankly, most of them are not worth recommending.

Microsoft XL

Probably the commonest way to collect data and perform simple manipulations is within a Microsoft XL spread-sheet. Consequently, the most obvious way to carry out statistical analyses of such data would seem to lie within XL itself. Let me give you my first piece of advice. Do not even consider it! The data analysis procedures within XL are rubbish – a very poor selection of procedures, badly implemented (apart from that, they are OK). If anybody in the Microsoft Corporation had a clue, they could have cleaned up in this area years ago. The opposition is pretty feeble and a decent statistical analysis package built into XL would have been unstoppable.

It is only at the most basic level that XL is of any real use (calculation of the mean, SD and SEM). It is therefore mentioned in some of the early chapters but not thereafter.

Other packages

A small survey by the author suggests that, in the subject areas mainly targetted by this book, only Minitab and SPSS are used with any frequency. Other packages seem to be restricted to very small numbers of departments. A decision was taken not to include blow-by-blow accounts of how to perform specific tests using any package, as this would excessively limit the book's audience. Instead, general comments are made about:

- entering data into packages;
- the information that will be required before any package can carry out the procedure;
- what to look for in the output that will be generated.

The last point is usually illustrated by generic output. This will not be in the same format as that from any specific package, but will present information that they should all provide.

Detailed instructions for Minitab and SPSS on the web site

As Minitab and SPSS clearly do have a significant user base, detailed instructions on how to use these packages to execute the procedures in this book, will be made available through the website associated with the book (www.staff.ljmu.ac.uk/phaprowe/pharmstats.htm). These cover how to:

- arrange the data for analysis;
- trigger the appropriate test;
- select appropriate options where relevant;
- find the essential parts of the output.

Part 1

Data types

1

Data types

This chapter will . . .

- Set out a system for describing different types of data
- Explain why we need to identify the type of data we are dealing with

1.1 Does it really matter?

To open a statistics book with a discussion of the way in which data can be categorized into different types probably sounds horribly academic. However, the first step in selecting a data handling technique is generally identifying what type of data we are dealing with. So, it may be dry, but it does have real consequences.

We will look at three types of data. All of these go under a variety of names. I have chosen names that seem to me to be the most self-explanatory, rather than sticking rigorously to any consistent system. The three terms that I will use are:

- interval scale – continuous measurement data;
- ordinal scale – ordered categorical data;
- nominal scale – categorical data.

Essential Statistics for the Pharmaceutical Sciences Philip Rowe
 ISBN 9780 470 03470 5 (HB) ISBN 9780 470 03468 2 (PB)

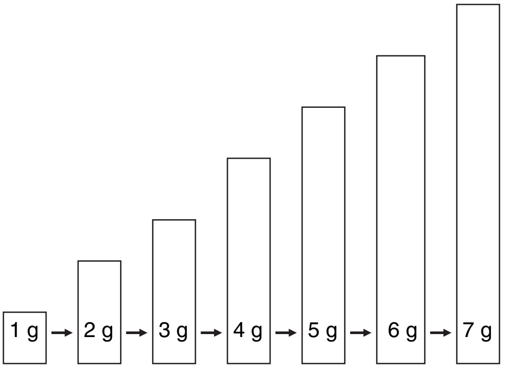

Figure 1.1 Interval scale data – a series of weights (1–7 g)

1.2 Interval scale data

The first two types of data that we will consider are both concerned with the measurement of some characteristic. 'Interval scale', or what is commonly called 'Continuous measurement', data include most of the information that would be generated in a laboratory. These include weights, lengths, timings, concentrations, pressures, etc. Imagine we had a series of objects weighing 1, 2, 3 up to 7 g as in Figure 1.1.

Now think about the differences in weights as we step from one object to the next. These steps, each of one unit along the scale, have the following characteristics:

1. *The steps are of an exactly defined size.* If you told somebody that you had a series of objects like those described above, he or she would know exactly how large the weight differences were as we progressed along the series.

2. *All the steps are of exactly the same size.* The weight difference between the 1 and 2 g objects is the same as the step from 2 to 3 g or from 6 to 7 g, and so on.

Because these measurements have constant sized steps (intervals), the measurement scale is described as a 'constant interval scale' and the data as 'interval scale'. Although the weights quoted in Figure 1.1 are exact integers, weights of 1.5 or 3.175 g are perfectly possible, so the measurement scale is said to be 'continuous'.

1.3 Ordinal scale data

Again measurement is involved, but the characteristic being assessed is often more subjective in nature. It is all well and good to measure nice neat objective things like blood pressure or temperature, but it is also a good idea to get the patient's angle on

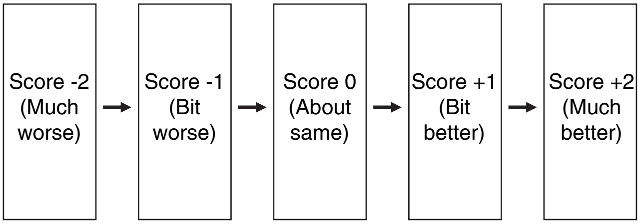

Figure 1.2 Ordinal scale data – scores for patient responses to treatment

how they feel about their treatment. The most obvious way to do this is as a score, of (say) -2 to $+2$ with the following equivalences:

-2 = markedly worse
-1 = a bit worse
0 = about the same
$+1$ = a bit better
$+2$ = markedly better

In this case (Figure 1.2), all we know is that, if one patient has a higher value than another, they are more satisfied with their outcome. However, we have no idea how much more satisfied he/she might be.

Since we have no idea how large the steps are between scores, we obviously cannot claim that all steps are of equal size. In fact, it is not even necessarily the case that the difference between scores of -2 and 0 is greater than that between $+1$ and $+2$. So, neither of the special characteristics of an constant interval scale apply to this data.

The name ‘ordinal’ reflects the fact that the various outcomes form an ordered sequence going from one extreme to its opposite. Such data are sometimes referred to as ‘ordered categorical’. In this case the data are usually discontinuous, individual cases being scored as -1, $+2$ etc., with no fractional values.

1.4 Nominal scale data

In this case there is no sense of measuring a characteristic. With this data we use a system of classifications, with no natural ordering. For example, one of the factors that might influence the effectiveness of treatment could be the specific manufacturer of a medical device. Therefore, all patients would be classified as users of ‘Smith’, ‘Jones’ or ‘Williams’ equipment. There is no natural sequence to these; they are just three different makes.

With ordinal data we did at least know that a case scored as (say) $+2$ is going to be more similar to one scored $+1$ than to one scored 0 or -1. However, with nominal data, we have no reason to expect Smith or Jones equipment to have any special degree of similarity. Indeed the sequence in which one would list them may be entirely arbitrary.

Quite commonly there are just two categories in use. Obvious cases are male/female, alive/dead or success/failure. In these cases, the data are described as 'dichotomous'.

Data types

- *Interval scale* – measurements with defined and constant intervals between successive values. Values are continuous.
- *Ordinal scale* – measurements using classifications with a natural sequence (lowest to highest), but with undefined intervals. Values are discontinuous.
- *Nominal scale* – classifications that form no natural sequence.

1.5 Structure of this book

The structure of this book is based upon the different data types. Chapters 2–14 all deal with the handling of continuous measurement data, with Chapters 15 and 16 focusing on categorical data, and then Chapter 17 covers ordered data.

1.6 Chapter summary

When selecting statistical procedures, a vital first step is to identify the type of data that is being considered.

Data may be:

- *Interval scale* – measurements on a scale with defined and constant intervals. Data are continuous.
- *Ordinal scale* – measurements on a scale without defined intervals. Data are discontinuous.
- *Nominal scale* – classifications that form no natural sequence.

Part 2

Interval-scale data