



WILEY

PRACTICAL  
STATISTICS  
for  
ENVIRONMENTAL  
and  
BIOLOGICAL  
SCIENTISTS

JOHN TOWNEND



# Practical Statistics for Environmental and Biological Scientists



# Practical Statistics for Environmental and Biological Scientists

**John Townend**

*University of Aberdeen, UK*



JOHN WILEY & SONS, LTD

Copyright © 2002

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester,  
West Sussex PO19 8SQ, England

Telephone (+44) 1243 779777

Email (for orders and customer service enquiries): [cs-books@wiley.co.uk](mailto:cs-books@wiley.co.uk)

Visit our Home Page on [www.wileyeurope.com](http://www.wileyeurope.com) or [www.wiley.co.uk](http://www.wiley.co.uk)

Reprinted with corrections March 2003.

Reprinted July 2004, April 2005, April 2006, April 2007, October 2008, April 2009

All Rights Reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except under the terms of the Copyright, Designs and Patents Act 1988 or under the terms of a licence issued by the Copyright Licensing Agency Ltd, 90 Tottenham Court Road, London W1T 4LP, UK, without the permission in writing of the Publisher. Requests to the Publisher should be addressed to the Permissions Department, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, or emailed to [permreq@wiley.co.uk](mailto:permreq@wiley.co.uk), or faxed to (+44) 1243 770571.

This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the Publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

#### ***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, 22 Worcester Road, Etobicoke, Ontario, Canada M9W 1L1

#### ***British Library Cataloguing in Publication Data***

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

ISBN 13: 978-0-471-49664-9 (HB)

ISBN 13: 978-0-471-49665-6 (PB)

Typeset in 10.5 / 13pt Times by Vision Typesetting Manchester

# Contents

*Preface*

ix

<b>PART I STATISTICS BASICS</b>	<b>1</b>
<b>1 Introduction</b>	<b>3</b>
1.1 Do you need statistics?	3
1.2 What is statistics?	4
1.3 Some important lessons I have learnt	5
1.4 Statistics is getting easier	6
1.5 Integrity in statistics	7
1.6 About this book	8
<b>2 A Brief Tutorial on Statistics</b>	<b>9</b>
2.1 Introduction	9
2.2 Variability	9
2.3 Samples and populations	10
2.4 Summary statistics	11
2.5 The basis of statistical tests	19
2.6 Limitations of statistical tests	24
<b>3 Before You Start</b>	<b>27</b>
3.1 Introduction	27
3.2 What statistical methods are available?	28
3.3 Surveys and experiments	33
3.4 Designing experiments and surveys – preliminaries	35
3.5 Summary	43
<b>4 Designing an Experiment or Survey</b>	<b>45</b>
4.1 Introduction	45
4.2 Sample size	45
4.3 Sampling	50
4.4 Experimental design	56
4.5 Further reading	60
<b>5 Exploratory Data Analysis and Data Presentation</b>	<b>63</b>
5.1 Introduction	63
5.2 Column graphs	65
5.3 Line graphs	67
5.4 Scatter graphs	69

5.5	General points about graphs	71
5.6	Tables	73
5.7	Standard errors and error bars	74
<b>6</b>	<b>Common Assumptions or Requirements of Data for Statistical Tests</b>	<b>77</b>
6.1	Introduction	77
6.2	Common assumptions	81
6.3	Transforming data	84
<b>PART II</b>	<b>STATISTICAL METHODS</b>	<b>91</b>
<b>7</b>	<b><i>t</i>-tests and <i>F</i>-tests</b>	<b>93</b>
7.1	Introduction	93
7.2	Limitations and assumptions	94
7.3	<i>t</i> -tests	95
7.4	<i>F</i> -test	103
7.5	Further reading	105
<b>8</b>	<b>Analysis of Variance</b>	<b>107</b>
8.1	Introduction	107
8.2	Limitations and assumptions	109
8.3	One-way ANOVA	111
8.4	Multiway ANOVA	119
8.5	Further reading	127
<b>9</b>	<b>Correlation and Regression</b>	<b>129</b>
9.1	Introduction	129
9.2	Limitations and assumptions	130
9.3	Pearson's product moment correlation	131
9.4	Simple linear regression	135
9.5	Correlation or regression?	142
9.6	Multiple linear regression	143
9.7	Comparing two lines	146
9.8	Fitting curves	148
9.9	Further reading	151
<b>10</b>	<b>Multivariate ANOVA</b>	<b>153</b>
10.1	Introduction	153
10.2	Limitations and assumptions	154
10.3	Null hypothesis	156
10.4	Description of the test	156
10.5	Interpreting the results	158
10.6	Further reading	161
<b>11</b>	<b>Repeated Measures</b>	<b>163</b>
11.1	Introduction	163
11.2	Methods for analysing repeated measures data	166







# Preface

Statistics wasn't forced upon the environmental and biological sciences; it has been absorbed into their practice because it was realized that it had something to offer. Statistical methods provide us with ways of summarizing our data, objective methods to decide how much confidence we can place in experimental results, and ways of uncovering patterns that are initially masked by the complexity of a dataset. Also, if we carry out scientific investigations according to our instincts, there is a risk that we will bias the results by overlooking some important factor or through our desire to get a particular result. By carefully following accepted statistical procedures we can avoid these problems and, just as importantly, we will be seen to have avoided them, so our results will be more readily accepted by others.

Statistics is also a useful means of communication. For example, a researcher might state that 'the molluscs had a mean shell length of  $12.2 \text{ mm} \pm 1.6 \text{ mm}$  (standard error)', or report that 'ANOVA showed significant differences between nitrogen contents in different groups of plants ( $P = 0.02$ )'. These are succinct ways of explaining a great deal of detail about how studies have been carried out and what can be concluded from them. Of course, they are only really a useful means of communication if you understand what the terms mean. Like it or not, though, they are widely used, so whether you intend to use statistics yourself or just read about others' research, it will still be a great help to know something about it.

While teaching statistics in a university I found that, for the most part, the statistical methods required by both environmental and biological scientists were the same. Indeed this might be expected, because much of the science is common to both as well. I also found that requirements were very similar at all levels from undergraduate to experienced professional. Really there is seldom any necessity to use complex statistical methods to do world-class research in environmental and biological sciences. Those who are able to identify the key, simple questions to ask are likely to enjoy the greatest success. So it is that I have tried to put together a book that addresses as many of the most common needs as possible.

The choice of content is based on the questions I have most frequently been asked and the explanations that seemed to work best. Memorizing formulae will be of very little practical use to you, except perhaps to pass an exam; most calculations can be carried out by computer these days. However, computers do

not generally tell you whether you are carrying out the right calculations or exactly what you can conclude from the results. Here textbooks still have a part to play. In this book I try to unlock many of the codes commonly used to present scientific information and to provide you with the tools you need to be an effective user of statistics yourself. I wholeheartedly hope that it will provide you with the information you need.

## **PART I**

# **STATISTICS BASICS**

Chapters 1 to 6 introduce the ideas behind statistical methods and how practical studies should be set up to use them. They aim to give the required background for using the methods in Part II. Readers who are new to statistics or in need of a short refresher might find it useful to read this part in its entirety.



---

# 1

## Introduction


---

If your first love was statistics, you probably wouldn't be studying or working in environmental or biological sciences. I am starting from this premise.

### 1.1 Do you need statistics?

Somebody who is trying to sell you a statistics textbook is probably not the best person to ask whether you need statistics. Maybe you have opened this book because you have an immediate need for these techniques or because you have to study the subject as part of a course. In this case the answer for you is clearly yes, you need statistics. Otherwise, if you want to know whether statistics is really relevant to you, ask people who have been successful in your chosen area – academics, researchers or people doing the kind of job you want to do in the future.

Some use it more than others, and certainly you will find some very successful people who are not confident with statistics and possibly dislike any involvement with it. I don't believe being a brilliant statistician is a necessary condition for being a brilliant biologist or environmental scientist. However, you will



## 1.2 What is statistics?

Football scores, unemployment rates and lengths of hospital waiting lists are statistics, but not what we commonly think of as being included in the subject of statistics. An interesting definition I heard recently was that statistics is ‘that part of a degree which causes a sinking feeling in your stomach’. I don’t have an all-encompassing definition myself, but it will be helpful if you can keep in mind that more or less everything in this book is concerned with trying to draw conclusions about very large groups of individuals (animate or inanimate) when we can only study small samples of them. The fact that we have to draw conclusions about large groups by studying only small samples is the main reason that we use statistics in environmental and biological science.

Supposing we select a small sample of individuals on which to carry out a study. The questions we are trying to answer usually boil down to these two:

- If I assume that the sample of individuals I have studied is representative of the group they come from, what can I tell about the group as a whole?
- How confident can I be that the sample of individuals I have studied was like the group as a whole?

These questions are central to the kind of statistical methods described in this book and to most of those commonly used in practical environmental or biological science. We are usually interested in a very large group of individuals (e.g. bacteria in soil, ozone concentrations in the air at some location which change moment by moment, or the yield of wheat plants given a particular fertilizer treatment) but limited to studying a small number of them because of time or resources.

Fortunately, if we select a sample of individuals in an appropriate way and study them, we *can* usually get a very good idea about the rest of the group. In fact, using small, representative samples is an excellent way to study large groups and is the basis of most scientific research. Once we have collected our data, our *best estimate* always has to be that the group as a whole was just like the sample we studied; what other option do we have? But in any scientific study, we cannot just assume this has to be correct, we also need to use our data to say how confident we can be that this is true. This is where statistics usually comes in.

Almost all experimental results are as described above. They state what is the case in a small sample that was studied, and *how likely* it is to be true of the group it was taken from. Elementary textbooks often quote results leaving out any indication of how much confidence we can place in them for the sake of clarity. However, most of the results they quote originally come from papers



published in scientific journals. If you look at the results presented in a scientific journal, you will see statements like:

Big gnomes catch more fish than little gnomes ( $P = 0.04$ )

The study would have been carried out using *samples* of big gnomes and small gnomes and the statement is really shorthand for:

- In our samples, on average, big gnomes caught more fish than little gnomes, so we expect that big gnomes in general catch more fish than little gnomes.
- Based on the evidence of our samples, we can really only be 96% confident that big gnomes in general do catch more fish than little gnomes.

You can see that the second, qualifying, statement (which comes from the  $P = 0.04$ ) is really quite important to understanding what the researchers have actually shown. It is not as clear-cut proof as you might otherwise think.

We will look in more detail at how to interpret the various forms of shorthand as we go through the different statistical techniques, but notice that when the result is stated in full we have (i) a result for the whole group of interest *assuming* that the samples studied were representative, and (ii) a measure of *confidence* that the samples studied actually were representative of the rest of the groups. This point is easy to lose sight of when we start to look at different techniques.

Textbooks tend to emphasize *differences* between statistical techniques so that you can see when to use each. However, these same ideas lie behind nearly all of them. Statistical methods, in a wide variety of disguises, aim to quantify both the effects we are studying (i.e. what the samples showed), and the confidence we can have that what we observed in our samples would also hold for the rest of the groups they were taken from. If you can keep this fact in mind, you already understand the most important point you need to know about statistics.

### 1.3 Some important lessons I have learnt

Statistics as a science in its own right can be very complicated. The statistics you need to be a good environmental scientist or biologist is only a small and fairly straightforward subset of this. Even a general understanding of the basic ideas will be a great asset when you come to interpret other people's experimental results. When you know some of the shorthand, like the example of the gnomes, you will see that very many scientific 'facts' are not as clear-cut and certain as we often imagine. Understanding just this already gives you statistical and

scientific skills beyond those of the general public. You will quickly learn to be more discerning about what scientific ‘facts’ you really believe.

There is no denying that a skilled statistician would have methods in his or her armoury beyond those I have included in this book. There are not statistical techniques available for *every* eventuality, but there are techniques for a good many of them. However, it takes rather a long time to learn about them all and probably you want to get on with some environmental or biological science too. I have therefore selected in this book a range of techniques that I consider most relevant and useful, and I believe these are sufficient to allow you to conduct most types of environmental or biological study *with a little careful planning*. Now here’s the bit that a lot of people find difficult to grasp. The thing that separates competent environmental scientists and biologists from incompetent ones, in terms of statistical skills, is not numeracy, but *careful planning*. The chances are that a computer will do all of your calculations for you.

By the time you sit down at the keyboard with your data you will have already made most of the mistakes you are likely to make. Just when you think you are about to *start* the statistical part of your project, your part in the statistics is really coming to an end. If you have planned carefully, formed a clear idea of what you are investigating, followed the layout of appropriate examples from this or other books, and carried out your survey or experiment accordingly, the analysis and interpretation will be plain sailing. Please don’t leave all thought of statistical analysis to the point where you sit down with your data already in hand. You would be unlikely to find the analysis plain sailing then. This is an important lesson I have learnt.

## 1.4 Statistics is getting easier

Until the 1980s most statistical calculations were done using a pocket calculator or by hand. Nowadays almost all calculations are carried out by computer. We need only know which test to use and how to enter the data in order to carry it out. I have heard concerns that many students nowadays just quote the output without understanding it. This is probably true, but it was always thus. As far as I can see, the only difference with precomputer days is that then you would spend two hours struggling with the calculations so there was a feeling you had earned the right to give the result. I don’t believe the average user of statistics either knew or cared what the calculations were actually doing any more than they do now.

Although I do not think that as *users* of statistics we need to do the calculations ourselves, we do lose a lot if we take the results without understanding anything about the methods. Until recently it was *necessary* to teach the calculations behind statistics because without them you could not use statistics, whether you understood them or not. To someone who is comfortable with

mathematical concepts, the formulae are also a satisfactory explanation of what is going on, so teachers often believed they had covered method and understanding at the same time.

An aunt of mine used to say, ‘There are liars, damn liars and sadistics. Most of the liars and damn liars go into law or advertising so don’t bother us much, but most of the sadistics teach numeracy skills. That’s why maths and statistics are hard.’ It is my belief that because statistics has traditionally been taught as a mathematical skill, although most students got by with the methods, very few picked up the understanding along the way. There is a great challenge here for teachers of statistics. Rather than seeing the removal of the calculations as a sad loss to understanding, we should take advantage of this to try to make the meaning and value of statistics more accessible to all.

## 1.5 Integrity in statistics

Scientific research *relies* on the integrity of the people conducting the research. Most of the time, we just have to believe that researchers have been honest in their work as there is no way to tell if results have been made up. In fact, in my experience very few people do lie about the actual values they have collected, even if they are disappointing. Most scientists, I think, have a fairly strong sense of conscience. We also need to have this attitude to carrying out an appropriate statistical analysis. Some kinds of analysis are easier to do than others and some may appear to give us the result we want whilst others do not. However, just because it is *possible* to use one statistical technique does not necessarily mean it is valid. Usually it is necessary to make certain checks on the data to discover whether a particular method can be applied validly (Chapter 6). Unfortunately, this can sometimes lead us to have to do more work, so there is a temptation to skip this stage.

The reader of our work, of course, has to assume that we *have* done the appropriate checks and, if necessary, carried out the extra work. Otherwise we should add the qualifying statement ‘assuming the test was valid in this case’, but then who would take our results seriously? If we just present results without checking the validity of using our chosen statistical method, we are deliberately deceiving the reader. If you value the integrity of your work, therefore, checking the validity of applying particular statistical methods must be seen as part of the normal process of statistical analysis. The checks required are described in the ‘Limitations and assumptions’ sections preceding each of the methods described in this book.

## 1.6 About this book

I have tried as far as possible to avoid mathematical descriptions of the techniques. There are a few simple formulae which readers might find occasion to use because they are not covered by some of the common statistical programs. I have included these in boxes; you can skip them if you want. I have also included some formulae in Appendix A, principally because they might be needed by some people for examinations in statistics. Mainly I have tried here to describe some of the range of techniques available, when you can use them, how to use them, and what the results are telling you; I have assumed that you will use a computer to do most of the calculations.

Competence and confidence in statistics will be an asset to you as an environmental or biological scientist, but at the same time it is only one of many things that will make you a good environmental or biological scientist. You only have so much time available, and to suggest you study the detail of statistics may not be the best use of it. With this in mind I have tried to include only techniques and a level of detail that I think will be genuinely useful to those studying or working in environmental or biological sciences.

There are different schools of thought about whether or not one should illustrate statistics with real experimental data. My own thoughts on this are that it is best to use simple datasets to demonstrate the techniques. It is not possible to cover all the eventualities that will arise in real-life results. However, provided you understand clearly what is required, you will be in a strong position to decide how to collect and handle your own data. All of the datasets in this book are therefore invented to demonstrate particular points.

The book is divided into two parts. Chapters 1 to 6 cover some basic statistical ideas and are intended to give you the necessary background for any of the statistical techniques in later chapters. Chapters 7 to 15 are more of a reference section with different statistical tests or methods described in each chapter. Guidance on which test to use in a particular situation is given in Section 3.2 and in the decision chart in Appendix D.

I have also included some pointers to more advanced techniques that readers might find useful in the further reading sections at the end of some of the chapters. If you have a computer package available to carry these out, understanding the details of the calculations need not necessarily be a problem to you. Nevertheless, before going ahead and using any of them it is important to familiarize yourself with what the tests are actually testing, and the assumptions and limitations they have about the types of data they are suitable for. In general, I have not specified particular texts to consult because these techniques are widely covered in many of the more in-depth statistical textbooks, and probably most of these would give you similar information.

---

# 2

## A Brief Tutorial on Statistics

---

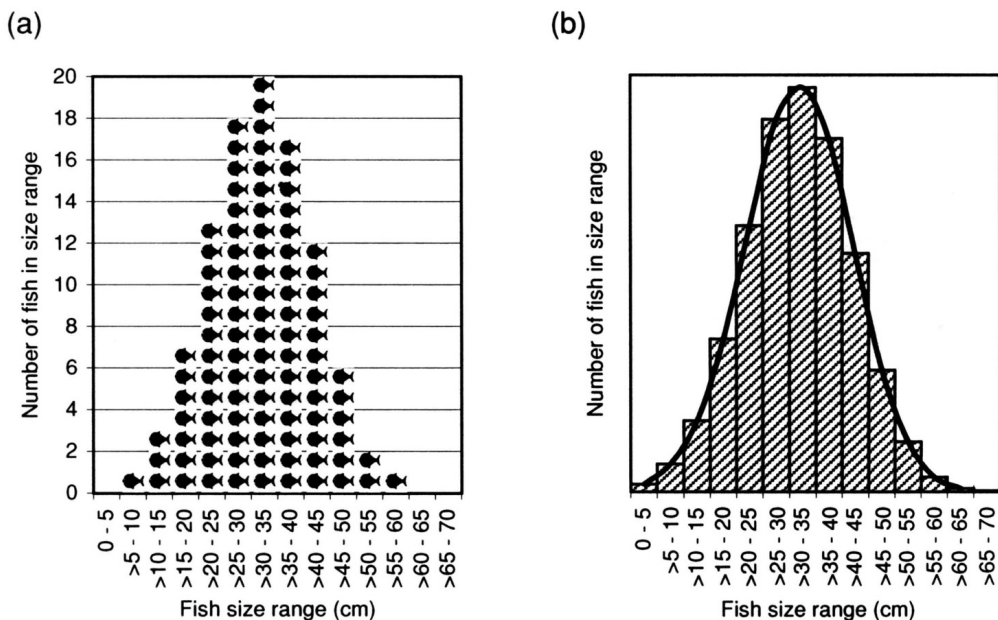
### 2.1 Introduction

From Chapter 3 onwards I describe a range of statistical tests and methods, and how to design experiments or surveys that make use of them. If you are studying this subject for the first time, you will probably find it difficult to retain all this information in your head. For the most part, this is not a problem. You can refer back to the book when you need to. However, there are some basic ideas behind all statistical methods and if you can keep these in mind, they will help you to make sense of statistical methods in general. These basic ideas are the subject of this chapter.

### 2.2 Variability

Think of a group you might want to study, e.g. the lengths of fish in a large lake. If all of these fish were the same length, you would only need to measure one. You can probably accept that they are not all the same length, just as people are not all the same height, not all volcanic lava flows are the same temperature, and not all carrots have the same sugar content. In fact, most characteristics we might want to study vary between individuals. If we measured the lengths of 100 fish, we could plot them on a graph as in Figure 2.1(a).

To understand this graph, think how we would add the extra point if we measured another fish to be 42 cm. It would appear as an additional fish in the column labelled > 40–45 cm. The graph tells us that most of the fish were about the same length, and gives us a picture of how widely spread the individuals' lengths were around this. We can see that only a few fish were greater than 50 cm or less than 15 cm. Figure 2.1(b) shows the more usual ways of presenting this kind of data.



**Figure 2.1** (a) Numbers of fish out of a sample of 100 falling into different size ranges. Note >10–15 means fish more than 10 cm long, up to and including 15 cm. (b) When larger numbers of measurements are involved, it becomes inconvenient to represent each individual, so a column graph or a line can be used to show the shape of the distribution

The graphs in Figure 2.1 are called *frequency distributions*. For some things we might measure, we would find different distributions such as a lot of low values, some high values, and a few extremely high values. These result in different shapes of graph (Sections 6.1 and 6.2). However, it turns out that if we measure a set of naturally occurring lengths, concentrations, times, temperatures, or whatever, and plot their distribution, very often we do get a diagram with a shape similar to those in Figure 2.1. This shape is called a *Normal distribution*. Statisticians have derived a mathematical formula which, when plotted on a graph, has the same shape. Being able to describe the distribution of individual measurements using a mathematical formula turns out to be very useful because, from only a few actual measurements, we can estimate what other members of the population are likely to be like. This idea is the basis of many statistical methods.

### 2.3 Samples and populations

As discussed in Chapter 1, practical considerations almost always dictate that we study any group we are interested in by making measurements or observations on a relatively small *sample* of individuals. We call the group we are

actually interested in the *population*. A population in the statistical sense is fairly close to the common meaning of the word, but can refer to things other than people, and usually to some particular characteristic of them. Here are some examples of statistical populations:

- The lengths of blue whales in the Arctic Ocean
- All momentary light intensities at some point in a forest
- Root lengths of rice plants of a particular variety grown under a specific set of conditions

In the first example, the population is real but we are unlikely to be able to study all of the whales in practice. Populations in the statistical sense, however, need not be finite, or even exist in real life. In the second example, the light intensity could be measured at any moment, but the number of moments is infinite, so we could never obtain measurements at every moment. In the third example, the population is just conceptual. We really want to know about how rice plants of this variety *in general* would grow under these conditions but we would have to infer this by growing a limited number of rice plants under the specified conditions. Although the few plants in our sample may be the only rice plants ever to be grown in these conditions, we still consider them to be a sample representing rice plants of this variety in general growing in these conditions.

## 2.4 Summary statistics

It is often useful to be able to characterize a population in terms of a few well-chosen statistics. These allow us to summarize possibly large numbers of measurements in order to present results and also to compare populations with one another.

### Mean, variance, standard deviation and coefficient of variation

If we want to describe a population it may sometimes be useful to present a frequency distribution like those in Figure 2.1, but this is usually more information than is needed. Two items are often sufficient:

- A measure which tells us what a ‘typical’ member of the population is like
- A measure which tells us about how spread out the other members of the population are around this ‘typical’ member

To represent a ‘typical’ member of the population, we usually use the *mean* (all

**Box 2.1   Standard deviation and variance**

One way to characterize how spread out the values in a sample are would be to calculate the difference between each measurement and the sample mean, and then to calculate the mean of these differences.

Here's an example:

Sample value	8	12	10	7	7	11	8	mean = 9.0
Difference from mean	1	3	1	2	2	2	1	mean of differences = $12/7 = 1.7$

Had statistics been developed after computers became readily available, this might be the measure of spread we commonly use; indeed some recently developed statistical methods do use this. However, when calculations were done by hand, it was found to be more convenient to use the mean of the *squares* of the differences as a measure of spread; there are mathematical shortcuts to getting the result in this case which are useful when you have a large sample. Since a great deal of statistical theory and tests built up around this, we still use it today.

For the above sample:

Sample value	8	12	10	7	7	11	8	mean = 9.0
Difference from mean	1	3	1	2	2	2	1	mean of differences = $12/7 = 1.7$
Square of difference from mean	1	9	1	4	4	4	1	mean of squared differences = $24/7 = 3.4$

When we take a random sample it may or may not include the largest and smallest values in the population, yet these would both contribute the largest *squares* of differences from the mean. Since they are not present in all samples, *on average* the mean of the squares of differences is less when it is calculated from a sample than if it was calculated for the population as a whole.

However, what we really want is to estimate the spread of values in the population, not in the sample itself, so we need to correct for this. This is done by modifying the above calculation so that we divide not by the number of values in our sample, but by *one less than the number of values in our sample*. You can see from the example below that this corrects things in the right direction. Some complex statistical theory shows that this simple modification corrects the value by the right amount.



For the above sample again:

Sample value	8	12	10	7	7	11	8	mean = 9.0
Difference from mean	1	3	1	2	2	2	1	mean of differences = $12/7 = 1.7$
Square of difference from mean	1	9	1	4	4	4	1	corrected mean = $24/6 = 4.0$

This figure – the corrected mean of the squared differences – is called the *variance* and is an unbiased *estimate* of the spread of values in the population, calculated from a sample.

Variance has units, e.g. if the measurements had been in grams (g), the variance would be in units of square grams ( $\text{g}^2$ ). The square root of the variance is called the *standard deviation*. The standard deviation in the above example is  $\sqrt{4.0} = 2.0$ , i.e. standard deviation is an alternative measure of the spread of values. Standard deviation has the same units as the actual measurements, e.g. if the measurements had been in grams, the standard deviation would also be in grams. The mathematical formulae for variance and standard deviation are given in Appendix A.

of our values added together then divided by the number of values). In common usage people often refer to this as the ‘average’ but the term ‘mean’ is preferred in technical writing.

To express how spread out the individual values in a population are, we usually use the *standard deviation* or *variance*; variance is simply the standard deviation squared (Box 2.1). In describing a population, we might therefore say, ‘The mean length of fish in the lake was 32 cm with a standard deviation of 10 cm.’ This tells us that most (approximately 68%) of the fish had lengths in the range 22–42 cm. Popular texts and the media often just give the mean with no measure of spread, but as scientists we should recognize that both measures are important. In a different lake the fish might have the same mean length but a very different spread of values. This might have important scientific implications. For example, if big fish eat little fish, the ecology of a lake with a wide range of sizes may be very different to that in a lake where the fish are all about the same size.

A further measure sometimes used to characterize the variability of a group is the *coefficient of variation* (CV). If we are told that the standard deviation of the lengths of ants is 3 mm, and the standard deviation of the lengths of dogs is 20 cm, we could correctly interpret this to mean that dogs are much more variable in their lengths. But we might also want to know which is more variable in relation to its size. If we divide the standard deviation by the mean, we get the coefficient of variation, i.e. the CV is the standard deviation relative

to the mean size of the individuals. For convenience, let's suppose the mean length of ants is 10 mm and the mean length of dogs is 100 cm. The CVs are therefore as follows:

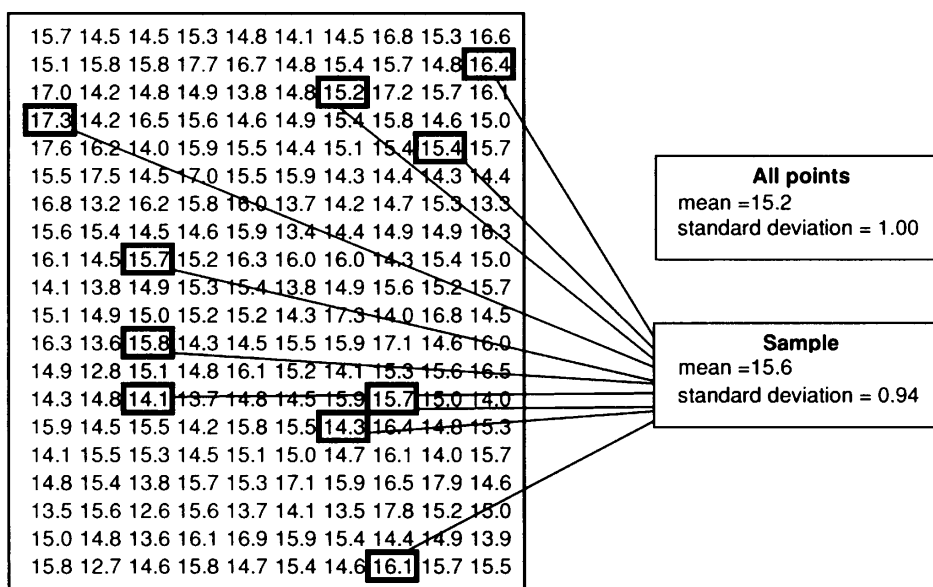
- For the ants  $3/10 = 0.3 = 30\%$
- For the dogs  $20/100 = 0.2 = 20\%$

Relative to their size, ants are more variable. The mean and standard deviation or CV are useful statistics to use to present the results of a survey.

### Standard error and 95% confidence interval

In experimental and survey work we are rarely interested in the samples we have actually studied. Our real interest is in the *populations* they come from. This is important to keep in mind, otherwise statistical tests make no sense.

Figure 2.2 shows the soil temperatures at 20 mm depth at different points in a field (the population). Suppose we want to know the overall mean temperature at 20 mm depth in this field – the *population mean*. The only way to find this out for sure would be to measure at every point, which would be impractical in a real field. What should we do? In most cases the best we can do is to use a *sample* of points and study them. We might make a start by measuring at 10 randomly



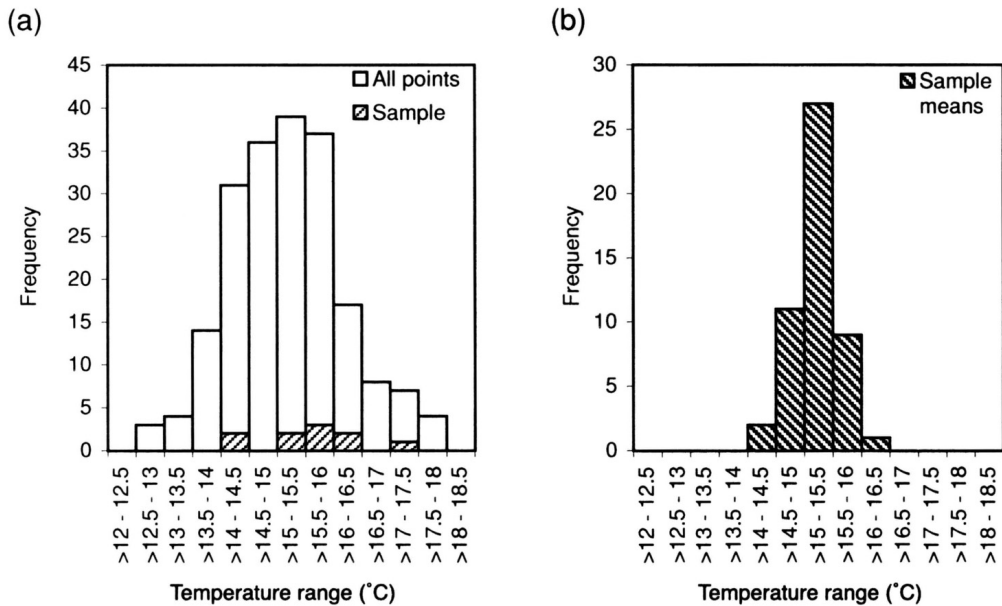
**Figure 2.2** Temperatures (°C) at 20 mm depth at 200 points in a field, and the mean and standard deviation of all the points and of a random sample of 10 points

selected points and calculating their mean temperature, the *sample mean* – we just add up the measurements and divide by 10 (Figure 2.2).

Of course, this still doesn't actually tell us the mean temperature for the whole field, it just tells us the mean temperature for those particular 10 points. Why is that any use? Figure 2.3(a) shows the distribution of temperatures in our sample, together with the distribution of temperatures for all the points given in Figure 2.2. As in this case, the distribution of values in a randomly selected sample is usually similar to that in the population as a whole, just with a lot fewer values. We can therefore say that a sample mean is probably a reasonable *estimate* of the population mean. What we need to do now is to try to give some measure of the 'margin of error' in this, i.e. to say, in any particular case, how reliable this estimate is likely to be.

So how can we tell what the margin of error is? Suppose I measured the temperature at another 10 points in the field at random and calculated their mean. I would probably get a slightly different value from the first sample. I could repeat this any number of times. Figure 2.3(b) shows the distribution of a series of 50 *sample means* obtained in this way. To understand this graph, imagine you took another random sample of 10 points and found their mean value was  $15.6^{\circ}\text{C}$ . This new point would be added to the top of the column labelled  $>15.5\text{--}16$ .

Notice in Figure 2.3(b) that although the different samples did not all give the



**Figure 2.3** (a) Distributions of the temperatures in the field and of the measurements in the sample. (b) Distribution of mean values from 50 random samples, each of 10 measurements

same mean, they are all quite close to one another, and are clustered around the population mean of 15.2°C. Sample means always cluster round the population mean like this. In this particular case, we can see that most of the sample means were within the range 14.5 to 16.0 (i.e. within 0.8 of the population mean). In other words, if we measured and calculated the mean of a randomly selected sample of 10 measurements from this field, we could be reasonably confident that the mean for the whole field would be within the range  $\pm 0.8$  of it. We have a measure of the 'margin of error' in this estimate, right here.

Although successive sample means are always clustered round the population mean like this, in other situations they may be clustered more or less tightly. To give a 'margin of error' round any particular sample mean, we need to know how tightly clustered a series of *similar* sample means would be in that situation. We could find this out by repeating every experiment 50 or so times but this would be impractical. However, it turns out that we can also estimate how tightly clustered a series of similar sample means would be from a single sample, by applying a simple formula (Box 2.2).

Looking at a graph like Figure 2.3(b) and saying that most means seem to be in a particular range is a bit subjective. Using the formulae in Box 2.2 to calculate the range of values in which most sample means would lie instead, achieves the same thing but also has the advantage that we get a more clearly defined and objective measure of the 'margin of error'. The most commonly used measure is called the *standard error* (Box 2.2). This is the range of values in which we can be approximately 68% confident that the true population mean lies. None of these 'margins of errors' *absolutely* defines the margin of error; the population mean might be a lot further out than this; we will never know for sure.

Therefore if we want to quote a 'margin of error' in which we can be more confident that the population mean really lies (in fact 95% confident), we can use the 95% *confidence interval* (95% CI). The 95% CI gives approximately twice as wide a range of values as the standard error (Box 2.2).

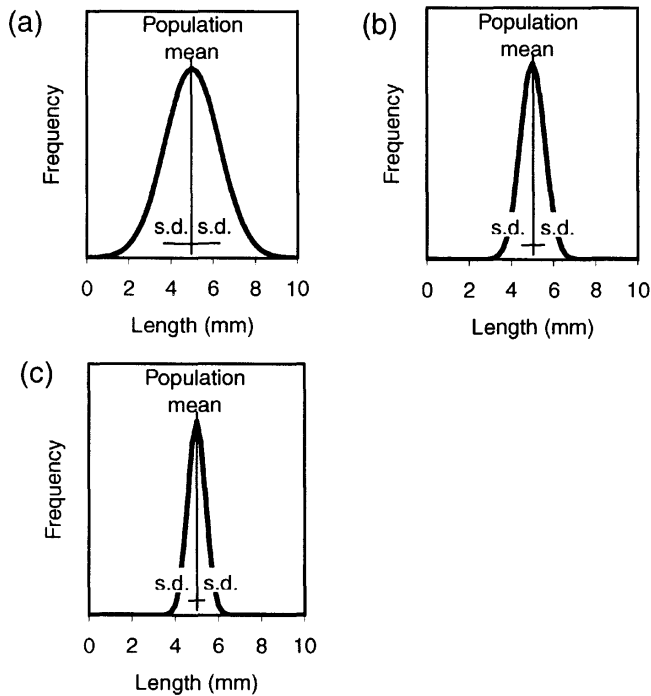
The result of a study may be stated as 12.5 mg  $\pm$  1.2 mg. This tells us that the mean value for the *sample* studied was 12.5 mg so this is our best estimate for the population mean. It also tells us that there is a considerable 'margin of error' in this estimate but we could at least be reasonably confident that the population mean was really somewhere in the range 11.3 to 13.7 mg. Either the standard error or the 95% CI can be shown in this way, so the text should make it clear which is given in any particular case.

A lot of people confuse the terms 'standard deviation' and 'standard error', presumably because they are both introduced about the same time in a course, both are new concepts, and both contain the word 'standard'. Concentrate on the words 'deviation' and 'error':

**Box 2.2** Standard error and 95% confidence interval

If we repeatedly measure randomly selected samples of, say, 5 individuals from a population, each time the sample mean will probably be slightly different. However, we would find that usually the sample mean was somewhere close to the mean of the population it came from. In fact, if we repeatedly measured samples of, say, 10 individuals, we would find that the means of these samples were even more closely clustered around the population mean (Figure 1). Remember how to interpret frequency distributions: the values are shown on the horizontal axis and the higher the line, the more often that value occurs.

We can calculate the standard deviation, i.e. the spread, of the *individuals* in the population as described in Box 2.1. It is also possible to calculate the standard deviation of the *sample means*. Figure 1 shows that the standard deviation of sample means is less for samples of size 10 than for samples of size 5. The standard deviation of the sample means is called the *standard error* (s.e.).



**Figure 1** Frequency distributions. (a) For a population. (b) For the means of a series of randomly selected samples of 5 individuals taken from the population in (a). (c) For the means of a series of randomly selected samples of 10 individuals taken from the population in (a). The standard deviation (s.d.) for each of the distributions is also shown