

Preface

I have written/edited several books on my own speciality, agricultural and horticultural entomology, but always at the request of a publisher or colleagues. This book is different in two ways. Firstly, it is a book I have positively wanted to write for many years and secondly, I am stepping out of my "comfort zone" in doing so.

The origins of this book stem from my appointment to the Horticulture Department of Reading University under Professor O. V. S. Heath, FRS. Professor Heath appreciated the importance of statistics and, at a time when there was no University-wide provision of statistics teaching, he taught the final year students himself. Following Heath's retirement, the course was given by his successor, Professor Peter Huxley, and I was asked to run the practical exercises which followed the lectures. You cannot teach what you do not understand yourself, but I tried nonetheless.

Eventually I took over the entire course. By then it was taught in the second year and in the third year the students went on to take a Faculty-wide course. I did not continue the lectures; the whole course was in the laboratory where I led the students (using pocket calculators) through the calculations in stages. The laboratory class environment involved continuous interaction with students in a way totally different from what happens in lectures, and it rapidly became clear to me that many biologists have their neurons wired up in a way that makes the traditional way of teaching statistics rather difficult for them.

What my students needed was confidence – confidence that statistical ideas and methods were not just theory, but actually worked with real biological data and, above all, had some basis in logic! As the years of teaching went on, I began to realize that the students regularly found the same steps a barrier to progress and damage to their confidence. Year after year I tried new ways to help them over these "crisis points"; eventually I succeeded with all of them, I am told.

The efficacy of my unusual teaching aids can actually be quantified. After having taken the Faculty course taught by professional statisticians, my students were formally examined together with cohorts of students from other departments in the Faculty (then of "Agriculture and Food") who had attended the same third year course in Applied Statistics. My students mostly (perhaps all but three per year out of some 20) appeared in the mark list as a distinct block right at the upper end, with percentage marks in the 70s, 80s, and even 90s. Although there may have also been one or two students from other courses with high marks, there then tended to be a gap until marks in the lower 60s appeared and began a continuum down to single figures.

I therefore feel confident that this book will be helpful to biologists with its mnemonics such as SqADS and "you go along the corridor before you go upstairs." Other things previously unheard of are the "lead line" and "supertotals" with their "subscripts" – yet all have been appreciated as most helpful by my students over the years. A riffle through the pages will amaze – where are the equations and algebraic symbols? They have to a large extent been replaced by numbers and words. The biologists I taught – and I don't think they were atypical – could work out what to do with a "45," but rarely what to do with an "x." Also, I have found that there are a number of statistical principles students easily forget, and then inevitably run into trouble with their calculations. These concepts are marked with

the symbol of a small elephant – "Never forget . . ." 

The book limits itself to the traditional foundations of parametric statistics: the *t*-test, analysis of variance, linear regression, and chi-square. However, the reader is guided as to where there are important extensions of these techniques. Finally the topic of nonparametric tests is introduced, but the calculation procedures are not explained. This is because the principles of the significant tests and therefore the calculations involved are outside the scope of this book; these calculations and the associated tables are readily available on the internet or in larger textbooks. However, the chapter does include one worked example of each test described as well as a check list of nonparametric methods linked to their parametric counterparts.

Many chapters end with an "executive summary" as a quick source for revision, and there are additional exercises to give the practice which is so essential to learning.

In order to minimize algebra, the calculations are explained with numerical examples. These, as well as the "spare-time activity" exercises, have come from many sources, and I regret the origins of many have become lost in the mists of time. Quite a number come from experiments carried out by horticulture students at Reading as part of their second year outdoor

practicals, and others have been totally fabricated in order to “work out” well. Others have had numbers or treatments changed better to fit what was needed. I can only apologize to anyone whose data I have used without due acknowledgment; failure to do so is not intentional. But please remember that data have often been fabricated or massaged – therefore do not rely on the results as scientific evidence for what they appear to show!

Today, computer programs take most of the sweat out of statistical procedures, and most biologists have access to professional statisticians. “Why bother to learn basic statistics?” is therefore a perfectly fair question, akin to “Why keep a dog and bark?” The question deserves an answer; to save repetition, my answer can be found towards the end of Chapter 1.

I am immensely grateful to the generations of Reading students who have challenged me to overcome their “hang-ups” and who have therefore contributed substantially to any success this book achieves. Also many postgraduate students as well as experienced visiting overseas scientists have encouraged me to turn my course into book form. My love and special thanks go to my wife Gillian who, with her own experience of biological statistics, has supported and encouraged me in writing this book; it is to her that I owe its imaginative title.

I am also most grateful for the very helpful, constructive and encouraging comments made by Professor Rob Marrs (School of Biological Sciences, University of Liverpool) and Dr Steve Coad (School of Mathematical Sciences, Queen Mary University of London). They both took the trouble to read the entire text at the request of the publisher and to do this they have given up a great deal of their time. Any misinterpretations of their advice or other errors are my fault and certainly not theirs.

Finally, I should like to thank Ward Cooper of Blackwells for having faith in this biologist, who is less terrified of statistics than he once was.

Helmut van Emden
May 2007