# Statistics Simplified

George F.K. Naylor
Laurie E. Enticknap

An Introductory Course
for Social Scientists and Others

# Statistics Simplified

## An Introduction for Social Scientists and Others

**George F.K. Naylor**
University of Queensland

**Laurie E. Enticknap**
University of Queensland

HBJ **Harcourt Brace Jovanovich Group**
(Australia) Pty Limited Sydney Melbourne

Printed in Australia

ISBN 0 7295 0209 0

To Rosemary and Geoffrey

# Foreword

*Statistics Simplified* has evolved in twenty years from a 20-page course for Adult Education students into an attractive and eminently understandable *personal address* to students of all Social Sciences. It is fitting that Dr. Naylor's Statistics course should have passed, on his retirement, to a former student, Mr. L.E. Enticknap, and that this joint publication should now be made available, with the willing cooperation of Mrs. Audrey Naylor, as a posthumous tribute to its originator.

Dr. Naylor's distinguished career began at the University of Sydney. He was awarded First Class Honours in Geology with the University Medal, and several prizes, including the coveted John Coutts Prize for the Science student graduating with the most distinguished academic record. He proceeded to the Master of Science degree, the Master of Arts degree in Psychology and the Diploma in Education. For ten years he was Assistant Director of the Australian Institute of Industrial Psychology, Sydney, and in World War II, as Squadron Leader, RAAF, he initiated and directed the selection procedures of that arm of the services, exercising his well-known flair for test construction and standardization.

In 1946 he joined the staff of the University of Queensland, where he was the first psychologist in Queensland to be awarded a Doctorate of Philosophy. During this period also, he was appointed to Fellowships in the Australian and the British Psychological Societies. On his retirement in 1976 his contribution to teaching and research was honoured by appointment as Honorary Research Consultant in the Department of Psychology. This post he held until his regrettable death on 16th July, 1980.

Those who knew Dr. Naylor will catch, in his illustrations and examples, glimpses of his familiar humour and some of his many and

varied interests. He viewed the whole universe of knowledge and skills as a personal challenge, and applied his experimental expertise equally to serious studies and to recreational pursuits. It is this intellectual approach which he endeavoured in his lectures and laboratory work to convey to his students. His teaching methods were original, attractive, lively and characteristically humourous — qualities which he had demonstrated in his choice of style for presentation of this subject. This intimate style will render *Statistics Simplified* readable, not only for the formal student, but also for the layman who wishes to know what statistics is all about.

Elsie Harwood
University of Queensland

# Preface

The primary purpose of *Statistics Simplified* is, as the name implies, to provide a simplified introduction to basic statistics. The style of writing is deliberately chosen to guide the individual student to an understanding of statistical concepts and of the various procedures, as they would be unfolded in a personal discourse.

The notes from which this book was developed were originally written by Dr. Naylor for students of psychology in introductory courses. They proved so popular that students proceeding to later courses at the University of Queensland frequently sought another copy to help them enter their more advanced work; while students from other departments contrived to obtain copies because they found them clearer and more understandable than the more sophisticated statistics texts.

The book contains many teaching devices which are not used in contemporary texts. Examples are: estimating correlation in the scattergram; establishing limits of ' about 2 ½ $\sigma$ ' to convey a picture of the distribution before discussing cut-off points as used in practice; introducing 'critical ratio' as a meaningful concept before discussing *t*; and other such techniques as promote understanding.

Additional details have been introduced in the Answers section for the benefit of interested students. This information has been made supplementary to the main text, to avoid complicating the on-going explanation for students who have difficulty with statistics or who require no more than a basic understanding of principles and methods.

It is for the benefit of students proceeding to later courses that three topics—Computational Formulae, Analysis of Variance and Nonparametric Statistics—have been added in Appendixes A, B and C respectively. While

the aim of the book is to make elementary concepts clear, these appendixes show statistics at work. They also complete the text for those whose research needs are fully satisfied by the use of these basic techniques.

Students who tend to be afraid of anything mathematical need not fear the simple concepts covered in this text. If, however, they do require more help—either for the purposes of this elementary course or because they are proceeding to a more advanced course—they would be advised to procure the little book, *Elementary Mathematics, an Introductory Course for Social Scientists and Others*, by Enticknap and Morgan. It provides a simple introductory approach to basic mathematics, and should allay anxiety about symbols.

I hope that *Statistics Simplified* will encourage many thinking people to gain a working knowledge of the subject. The charlatan is ever ready to misuse statistics to persuade the uninitiated.

L.E.E.

# Acknowledgements

Many people have contributed to the making of *Statistics Simplified* and to ensuring that it will live up to its title. I would like to thank them all, though I can name only a few—in particular, my wife Heather, without whose attention, help and devotion, *Statistics Simplified* could never have been published.

Since the sudden passing of Dr. Naylor, I have had constant co-operation from Mrs. Audrey Naylor, who has taken an interest in every step of the work.

My very special thanks go to Dr. Elsie Harwood who, having co-operated with Dr. Naylor in checking earlier editions, has spent many long hours reviewing and checking the clarity and accuracy of both draft material and the final manuscript. Her suggestions have improved the text immeasurably.

Thanks are also due to Mr. Richard S. Mortimer-Tanner for his help over many revisions in checking exercises and helping with the less glamorous side of the work. Marilyn Astle also contributed considerably in this respect, as well as providing many of the exercises and helping to improve the text. My sincere thanks go to her.

I also wish to thank Dr. Roderick Ashton, Head of the Department of Psychology, for his general interest and encouragement.

I am most grateful to Irene Hall and Sannie Pritchard, who spent so much time and gave so much care to the unenviable task of typing the final copy in its published form. In many places they have really accomplished the impossible. My thanks also go to my niece Elizabeth Metcalfe, who volunteered to type a draft and so assist in bringing the material into due form.

Finally, but by no means least, I would like to thank those of my students, as well as Dr. Naylor's, who, by their inspiration, suggestions and support, contributed so much to the success of *Statistics Simplified.*

L.E.E.

# Contents

# 1
# The Nature and Aims of Statistical Calculation

Students of any scientific discipline should have a clear concept of the nature and usefulness of statistical procedures. The social sciences are certainly no exception.

The connection between statistics and mathematics is widely recognized and those whose record in mathematics is somewhat unsatisfactory often approach statistics in a rather emotional manner which may be little short of an expectation of defeat. This unfortunate attitude is, however, entirely unnecessary and unjustified. For, while it is true that a competent professional statistician must possess much mathematical training and comprehension, the amount of basic mathematical skill required by the ordinary scientist to cope with routine applications of statistics to his studies is surprisingly small. What is needed is a capacity for clear logical thinking, often greatly assisted by an equal capacity for intelligent approximation. But surely these can be presupposed in students at tertiary level.

Statistical processes employ arithmetic, but the purpose of their use is specifically directed to making generalized as distinct from specific conclusions. Students at the elementary level will be saved a great deal of trouble if they can recognize from the start the difference in *aim* between a statistical calculation and a piece of arithmetic.

When you trundle your trolley to the check-out of the supermarket, a piece of arithmetic has to be accomplished. The prices of the individual items have to be added up to determine how much you will have to pay. This is usually done mechanically with the cash-register. The answer is quite definite - $4·84, $2·29 or whatever the case may be.

The money is paid and the incident is closed. The calculation has no ulterior purpose, it is *not* a statistical one. However, supposing you were to put all the cash dockets in your pocket and at the end of the week you had a collection of five of them, you would be in a position to do something really statistical. You could add up the totals on each day's docket and divide the result by the number of dockets. This would give the average amount spent each day, and in so far as you consider that the past week gave a fair sample of your buying habits, you could make a more or less reliable estimate of how much you are likely to spend in any given future period, or even how much you probably spent in any past period for which you had no exact record.

The daily average which you have calculated for your week's spending is not a real quantity, in that it is probably not exactly equal to the actual spending on any particular day, nor would the storekeeper be very satisfied if you were to offer him just this amount for your next day's purchases. Yet an average is a valuable statistical concept, because it is the best way of representing with a single figure in a general way the daily extent of your spending.

There are other statistical procedures which could be used with greater or less advantage for the same general purpose, but that is another story. What is important here is to recognize that a *statistical* calculation gives information that has a generalized (even if not quite exact) application. We are interested in the answer not for itself alone, but for what we can infer from it about other related circumstances.

It is easy to realize that the accuracy of such deductions will depend upon the extent to which the sample used for our statistical calculation was a fair, representative one. If we were to base our estimate of daily spending at the supermarket on dockets for a week in which we had done an unusual amount of entertaining, our average would be an over-estimate; whereas the figure for a week in which a member of the family had been absent would probably give an under-estimate.

In the social sciences we are constantly observing situations and carrying out experiments. Although we are interested in the outcome of these particular exercises, our real purpose is to infer what is *generally* likely to happen in circumstances of a similar nature. In other words, we are looking for laws or rules that govern human behaviour in a general sense. Hence we must endeavour to make statistical calculations from our actual observations.

We use the particular experiment as a *sample* upon which to base future expectations. Since we have no means of knowing just how really representative any particular sample is, we cannot be certain that our predictions or expectations based upon it will be completely accurate. Inferences based on statistical calculations, however accurately we do the arithmetic, are always liable to *statistical* error. However, using the mathematical theory of probability, it is usually possible to estimate the probable amount of error involved, so that at least we know within what limits the true value that we are seeking is likely to lie.

It is the task of the student to familiarize himself with the basic principles underlying the most important statistical procedures, including the estimation of the extent of errors involved. Only then will he be in a position to appreciate the general significance of work which he himself performs, or which is reported to him by other social scientists.

Like most other technical procedures, statistics may involve many delicate refinements and adjustments to meet special circumstances, but it is more important for the social scientist to be aware of the circumstances when some additional adjustment should be made than to be able to apply this for himself. The services of a mathematically trained statistician are usually available to assist him in this regard - either personally or through consultation of a detailed textbook. However, an understanding of the general aims and methods of statistics is quite possible without concern for the cunning devices which the more experienced mathematician may employ to shorten calculations or provide slightly greater degrees of accuracy.

In this course we shall be concerned more with basic principles than with sophisticated methods. There are many computational short-cuts, but we shall not hesitate to avoid these if the roundabout way is more easily understandable. The student who wishes to do so will have plenty of opportunity to absorb the technical detail at a later stage.

### Exercises on Chapter 1

Answer true or false to each of the following statements:

1. Statistical analysis and arithmetical calculation have the same aims.
2. For statistical analysis, the social scientist requires:
   (a) extensive mathematical training.
   (b) logical thinking.
   (c) capacity for approximation.

(d) more than average mathematical skill.

3. Any sample may be regarded as representative.
4. Your calculated average daily expenditure for a week
   (a) enables some prediction of future daily expenditure.
   (b) can be described as a statistical calculation.
   (c) can be described as a real quantity.
   (d) may be a fair sample of your spending habits.
5. Inferences based on statistical calculations, given accurate arithmetical work
   (a) are always liable to error.
   (b) can be checked for probable amount of error.
   (c) should be checked by using another sample.