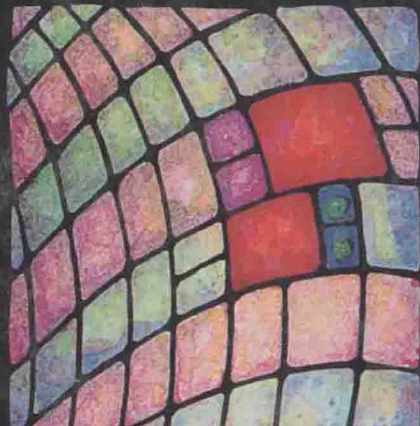
# BASIC STATISTICAL METHODS

## FIFTH EDITION



N. M. DOWNIE
ROBERT W. HEATH

FIFTH EDITION

# BASIC STATISTICAL METHODS

N. M. DOWNIE
Professor Emeritus of Psychological Sciences,
Purdue University

R. W. HEATH
Nomos Institute

1817
HARPER & ROW, PUBLISHERS, New York
Cambridge, Philadelphia, San Francisco,
London, Mexico City, São Paulo, Sydney

Sponsoring Editor: George A. Middendorf
Interior Design: Hudson River Studio
Production Manager: William Lane
Compositor: Kingsport Press
Printer and Binder: R. R. Donnelley & Sons Company
Cover Designer: James Laughlin

Basic Statistical Methods, Fifth Edition



**Library of Congress Cataloging in Publication Data**

Downie, N. M. (Norville Morgan), 1910–
Basic statistical methods.

Includes bibliographical references and index.
1. Social sciences—Statistical methods. 2. Statistics. I. Heath, R. W. II. Title.
HA29.D68 1983 519.5 83–213
ISBN 0–06–041728–5

Harper International Edition
35–02663

# PREFACE

This edition, like its four predecessors, is written to meet the need of the beginning student in the social sciences for a short, clear, elementary statistics book. We believe that such a book should treat the computation, interpretation, and application of commonly used statistics. No extensive attempt has been made to derive formulas or to involve statistical theory, since the mathematical background of the typical user of this book precludes effective presentation of these topics.

In this edition the first five chapters are concerned with the simpler descriptive statistics. These are followed by three chapters on regression and correlation. The next nine chapters consist of an introduction to statistical inference. The last chapter is an introduction to test theory and construction.

This book differs from the four previous editions mostly in that the statistical methodologies are adapted for use with the pocket calculator, this common tool of the modern student. We think that because these calculators are so universal, the beginning statistics course should be geared to them. The use of these calculators then makes it possible to eradicate the tedious hand methods using grouped data. Appendix 1 contains some practice exercises to help students familiarize themselves with their calculators and the processes used in statistics.

Many portions of the fifth edition have been rewritten, and new material has been added. Altogether, we have attempted to present an up-to-date elementary text for a rapidly developing field.

Problems designed to offer practice in the techniques discussed in each chapter appear throughout the book. The answers appear in Appendix O.

chapter appear throughout the book. The answers appear in Appendix O. At the request of many users of this textbook, the senior author has written a separate study guide for use with the text.

Appendixes A to M contain material that will become more and more useful as the student progresses through the book. Appendixes B through M are associated with the statistical concepts and tests introduced in the text.

Many sincere thanks are due to the various authors and publishers who gave us permission to use the material that appears in Appendixes A to M. Special acknowledgments appear at the end of each appendix.

We are especially indebted to the late Ronald A. Fisher of Cambridge, to Dr. Frank Yates of Rothamsted, and to Messrs. Oliver and Boyd, Ltd., of Edinburgh, for permission to reprint Appendixes C, D, and F from their book *Statistical Tables for Biological, Agricultural, and Medical Research.*

**N. M. Downie**
**R. W. Heath**

# CONTENTS

PART ONE

# DESCRIPTIVE STATISTICS

C H A P T E R 1

# Introduction

Students often wonder what statistics is and why they should bother to study the subject. For hundreds of years people have been collecting statistics. An early tribal chief, for instance, had so many armed warriors, so many horses, took a certain number of the enemy. Today we have vast quantities of data associated with sports, the stock market, traffic, law enforcement, and hundreds of other human activities. From one point of view, then, statistics may be considered collections of data associated with human enterprises. In a more limited sense, each individual is a statistic. From the viewpoint of a life insurance company, each of us is a statistic.

Statistics may also be considered to be a method that can be used to analyze data, that is, to organize and make sense out of a large amount of material. It is with this manipulation of data that we shall be concerned in this book. Statistical methodology may be looked upon as being of three types—descriptive, correlational, and inferential.

To illustrate the three types, suppose that we use the entering first-year class of a large university. Each student's folder will contain the scores of the admissions tests used by the university, such as the college boards, a high school transcript, results of a physical examination, and, perhaps, scores based upon interest and personality inventories. Taken as a whole, these folders present a mass of information about the class members. To learn about the whole class, all these data must be studied. Let us look at a few things that we can do with them. Suppose that we use only the scores on the verbal and mathematical parts of the Scholastic Aptitude Test. We might summarize all of these scores into two distributions. We might draw graphs that would show the differences in the scores of males and females, the differences among individuals in the several schools of the university, or the differences among students from varying backgrounds. Then we could find an average score on each of the two parts of the test for the whole group and for the different subgroups. If we are concerned with the relationship of individual scores to these averages, we can change these raw scores into another type of a more meaningful nature. *Centiles* or *standard scores* show how an individual in the group stands in reference to others. All of the foregoing operations are included in what we call *descriptive statistics* because they give us information, or describe, the sample we are studying.

We could also examine the scores on the verbal part of this test and see if they are related to scores on the mathematical part. This is called *computing a correlation coefficient.* We could also correlate these test scores with the first-semester grades made by these classmates, or with their high school ranks, or with the various scores on the interest and personality inventories. The results of correlational work are useful in making predictions of future behavior. If we know that a relationship exists between two variables, then scores on one may be used to predict scores on the other. In statistics this study of prediction is referred to as regression analysis. The results of correlational analysis are used to study the reliability and validity of educational and psychological tests. *Correlational analysis,* then, is a major part of statistical methodology.

Third we have *inferential statistics.* Usually when samples are studied, the investigator is interested in going beyond the sample and making an inference about the population from which the sample was drawn. Populations are frequently so large that the only way their characteristics will ever be known is through the study of samples drawn systematically from the population. It follows, then, that from measures of averages and variability based upon samples we make inferences about the size of the same traits in the population. The use of inferential statistics is basic to experimental research in all branches of science.

There are reasons for studying statistics other than knowing how to use the subject in a research task. A knowledge of statistics is basic to the intelligent reading of a research article or a modern text in science. Without the background one would get from a first course in statistics, these accounts of modern science are unintelligible. Statistics is also of use to the informed instructor in building and analyzing tests and in preparing grades. Statistics may also contribute to the general education of a consumer. Modern advertising makes all sorts of claims, often bolstering them with impressive statistics. The intelligent consumer looks critically at these claims and the statistics used to support them.

## A BRIEF HISTORY OF STATISTICS

Statistics has a long and venerable history. Perhaps the earliest use of statistics was when an ancient chief counted the number of the tribe's effective warriors or the number needed to defeat the enemy, or when the ruler figured how much might judiciously be collected in taxes. In later times, statistics were used to report death rates in the great London plague and in the study of natural resources. These uses of statistics, which encompass a broad field of activity referred to as "state arithmetic," are purely descriptive in nature.

In the seventeenth and eighteenth centuries mathematicians were asked by gamblers to develop principles that would improve the chances of winning at cards and dice. The two most noted mathematicians who became involved in this, the first major study of probability, were Bernoulli and DeMoivre. In the 1730s DeMoivre developed the equation for the normal curve. Important work on probability was conducted in the first two decades of the nineteenth century by two other mathematicians, LaPlace and Gauss. Their work was an application of probability principles to astronomy.

Through the eighteenth century statistics was mathematical, political, and governmental. In the early nineteenth century a famous Belgian statistician, Quetelet, applied statistics to investigations of social and educational problems. Walker (1929)[1] credits Quetelet with developing statistical theory as a general method of research applicable to any observational science. Beyond any doubt, the individual who had the greatest effect upon the introduction and use of statistics in the social sciences was Francis Galton. In the course of his long life he made notable contributions in the fields of heredity and eugenics, psychology, anthropometry, and statistics. Our present understanding of correlation, the measure of agreement between two variables, is credited to him. The mathematician Pearson collaborated with Galton in later years and was instrumental in developing many of the correlation and regression formulas that are in use today. Among Galton's contributions was the development of centiles or percentiles.

The famous American psychologist James McKeen Cattell studied in Europe in the 1880s and contacted Galton and other European statisticians. On his return to the United States he and his students, including E. L. Thorndike, began to apply statistical methods to psychological and educational problems. The influence of these men was great; in a few years theoretical and applied statistics courses were commonly taught in American universities.

In the twentieth century new techniques and methods were applied to the study of small samples. The major contributions in small-sample theory were made by the late R. A. Fisher, an English statistician. Although most of his methods were developed in an agricultural or biological setting, it was not long before social scientists recognized the utility of Fisher's methods and made use of his ideas. Today statistics is the major methodological tool of the research worker in the social sciences. The student interested in the history of statistics is referred to a brief but thorough article written by Dudycha and Dudycha (1972).

---

[1] Complete references are given at the back of the book.

## HOW VARIABLES ARE CONSIDERED IN STATISTICS

### Types of Measurements

We can classify data into two types: continuous and discontinuous, or discrete. Feet, pounds, minutes, and meters are examples of continuous data. With these we can make measurements of varying degrees of precision. For example, we can break meters into centimeters, centimeters into millimeters, and with intricate devices we can make measurements that are more and more precise. Such data can be considered as points on a line. The size and accuracy of the measurements that we can make along this line depend on the way that the measurements are made.

To illustrate, suppose that a boy is measured and found to be 150 centimeters tall. Does this mean that he is exactly 150 centimeters tall? Probably not, because in reading the measurement scale we merely read that number of centimeters to which the boy's height was closest. This 150 centimeters includes a segment of a line, that is, the segment extending from 149.5 to 150.5 centimeters. Similarly, a reading of 151 centimeters extends from 150.5 to 151.5. Each measurement on a scale of continuous data has a lower and upper limit, as shown in Figure 1.1.

Discontinuous or discrete data, on the other hand, are based upon measurements that can only be expressed in whole units. The counting of people, for example, can only occur in whole units, in contrast to measurements of length, which can be divided into smaller and smaller units. Other examples of discrete units are the number of words spelled correctly, the number of objects assembled, and the number of cars passing a point during a certain period of time. The student will note, however, that in statistical work most data tend to be treated as continuous, so we make such statements as: The typical graduate of college A has 2.8 children. The student should become accustomed to thinking of every number as having an upper and a lower limit.

### Measurement Scales

Stevens (1946, 1958, 1968) has written extensively on the types of measurements that are used in science. While not all statisticians agree with Stevens

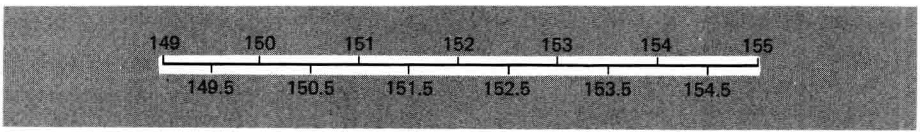



**FIGURE 1.1**
The upper and lower limits of continuous data.

on the types of statistics that should be used with the various kinds of measures in his classification, Stevens has devised a system that makes a logical approach to measurement. As he has pointed out, if there were no measurement there would be no statistics, and if measurements were accurate in all situations there would be a much lessened demand for statistics.

Stevens recognizes four types of scales: *nominal, ordinal, interval,* and *ratio. Nominal* scales are used as measures of identity. Numbers may serve as labels to identity items or classes. The numbers carried on the backs of athletes represent a nominal scale in its simplest form. Other examples of such scales are classifications of individuals into categories. For example, a sample of people being studied may be sorted into the following categories on the basis of religious preference: (1) Protestant; (2) Catholic; (3) Jewish; (4) other; and (5) none. Or they might be classified on the basis of sex, eye color, political party membership, urban–rural, and the like. Simple statistics are used with nominal data. For example, the number, proportion, or percentage of cases in each class or category may be determined.

When an *ordinal* scale is used in measurement, numbers reflect the rank order of the individuals or objects. Ordinal measures are arranged from the highest to the lowest or vice versa. The classical example of such a scale is the one used in evaluating the hardness of minerals. Hardness is defined as the degree of resistance to abrasion or scratching. On this scale 1 is characterized as being very soft and easily scratched, such as is the case with talc. At 10, the opposite end of the scale, is the diamond, which scratches all others and is itself scratched by none. Similarly, a group of men may be arranged by physical or mental traits. Ordinal measures reveal which person or object is larger or smaller, brighter or duller, harder or softer, etc., than the other. But such measures do not tell how much taller or how much heavier one is from the other. Statements such as "James is taller than John, who is taller than William, who is taller than Paul" can be made. Not much can be done with ordinal measures statistically except to determine the median and centiles and to compute rank correlation coefficients.

The third type of scale, the *interval* scale, provides numbers that reflect differences among items. With interval scales the measurement units are equal. Examples of such scales are the Fahrenheit and Celsius thermometers, time as reckoned on our calendar, and scores on intelligence tests. In the latter case we assume equal units of measurement. Many statistics are used with interval scales: arithmetic mean, standard deviation, and the product-moment correlation coefficient. Also, our most widely used statistical tests of significance, the $t$ test and the $F$ test, may be used with such data. Interval scales show that a person or item is so many units larger or smaller, heavier or lighter, brighter or duller, etc., than another.

The final and highest type of scale is the *ratio* scale. The basic difference between this type and the preceding one is that ratio scales have an absolute zero. It is true that interval scales (e.g., Fahrenheit and Celsius) also have zero points, but such points are arbitrarily chosen. Common ratio scales are measures of length, width, weight, capacity, loudness, and so on. In measuring temperature, the Kelvin scale, which has a zero point at −273°C where there is a complete absence of heat, is of this type. When a ratio scale is used, numbers reflect ratios among items, and data obtained with such scales may be subjected to the highest types of statistical treatments.

When data are in terms of feet, we can say that one length is twice or half that of another. When our measurements are on an interval scale, we cannot do this and make sense. For example, suppose that the maximum temperature today is 60°; the same day last year it was 30°. In this case we cannot state that it is twice as warm today as it was on the same date last year. What is the difference between these two conditions? When we were dealing with feet, we were using a measuring scale that was based upon an absolute zero; in the second case we are using a scale that started 32 degrees below the freezing point of water. When measurements are on a ratio scale, meaningful comparisons can be made. As a matter of fact, when data are of this type, all of the usual mathematical and statistical manipulations may be made. However, in actual practice many of our measurements are based upon interval scales and we apply practically all of our statistical techniques to these measurements.

What can we say about the measurements that we make in education, sociology, and psychology? First of all, we frequently assume that they have equal units of measurement. An inspection of certain of these, such as intelligence quotients, reveals that this assumption is not likely to be true. Furthermore, our scales do not possess an absolute zero. The physicist can describe absolute zero on his heat scale. It is not difficult to visualize zero inches, pounds, or meters. But what does zero IQ mean? Or what does it mean when a boy gets a score of zero on a geography test? Actually we do not know what these scores of zero mean. Then it follows that we have no basis for stating that a child with an IQ twice the size of the IQ of another child is twice as bright. Neither can we say that the child whose score on an arithmetic test is double that of another child knows twice as much arithmetic as the first child.

## SAMPLES AND POPULATIONS

It is important to distinguish between the terms *samples* and *populations.* Let us use an illustration to do this. Suppose that we are interested in the mental ability of children in the second grade. One way to investigate this is to give intelligence tests to second graders. We begin by obtaining permission to administer the test to one group of second graders. We compute the average, or mean, score for the group; this mean score is a statistic and gives us the average test score of our sample. Since there are so many second-grade pupils, we could continue this process for a long time by drawing sample after sample. If each of these samples is a *random sample* (random sampling procedures will be discussed in Chapter 9), we can combine all the sample averages, or means, to obtain a grand mean. This grand mean will be our best estimate of the average intelligence of all the second graders. That is, the average of all of the means of our samples is used to tell us something about the population value. All second graders in the United States make up the population, or universe, from which the various samples are drawn. Values that refer to populations are called *parameters;* values that refer to samples are called *statistics.* Populations, as the term is used in statistics, are arbitrarily defined groups. They need not be as large as the one used here as an illustration. We could define the 552 seniors in a certain school system as our population, and from this we could draw samples. One of the major aspects of statistical research is making inferences about population characteristics on the basis of one or more samples that have been studied.

## ON THE USE OF NUMBERS

Because statistics is a branch of mathematics and because many students using this book are far removed from a course in mathematics, it is suggested that the student review the basic mathematical and algebraic processes that are set forth in Appendix A. It is also recommended that the student buy a pocket calculator, at least an inexpensive one, to facilitate the solution of problems. In the past considerable time was spent teaching students how to take the square root of a number. Since this is done so rapidly and efficiently with a pocket calculator, it is a waste of time to do it in any other fashion. In the following paragraphs a few of the problems and conventions in the use of numbers are presented.

### Rounding Numbers

In rounding numbers to the nearest whole number or to the nearest decimal place, we proceed as follows:

| | |
|---|---|
| To the nearest whole number | $7.2 = 7$ |
| | $7.8 = 8$ |
| To the nearest tenth | $7.17 = 7.2$ |
| | $7.11 = 7.1$ |
| | $.09 = .1$ |
| To the nearest hundredth | $7.177 = 7.18$ |
| | $.674 = .67$ |
| | $1.098 = 1.10$ |

The general rule is that if the last digit is less than 5, it is dropped; if the last digit is more than 5, the preceding digit is raised to the next higher digit. The only complication arises when numbers end in 5. There is a general rule for this case. When the digit preceding the 5 is an odd number, this digit is raised to the next higher one; when it is an even number, the 5 is dropped. The following examples illustrate this rule:

$$8.875 = 8.88 \qquad 5.25 = 5.2$$

$$8.05 = 8.0 \qquad 66.975 = 66.98$$

Situations like the following arise:

$$\frac{37}{52}(3)$$

There are two ways of simplifying this. The first is to divide 37 by 52 and then multiply the quotient by 3. Or the numerator, 37, could be multiplied by 3 and the product then divided by 52 or multiplied by the reciprocal of 52. The second method is preferred because only one rounding operation is necessary.

Sometimes students get into trouble as a result of rounding numbers too freely in their problems. Suppose that we have an operation that consists of six distinct steps. At the conclusion of the computations for each step, the student rounds the results. A series of a half dozen such roundings in the course of the solution of a problem causes inaccuracies to enter the work. If we are going to express our answer to the nearest tenth, a good rule is to carry all operations through in terms of hundredths and round to the nearest tenth in the last step.

## Significant Digits

In recording numbers the question frequently arises as to how many digits we should have in our answers. As a general rule, the answer should have only one or two digits more than exist in the raw data. For example, if we have a series of test scores, each of which contains two digits, then ordinarily we would have no more than three digits in the average or mean that we compute from these data. There is nothing to be gained in computing these averages to five or six decimal places. No meaningful accuracy is obtained from these large decimals. As a matter of fact, such large decimals mean nothing when computed on the basis of two-place numbers. A good rule is to have one more significant digit in the answer than was present in the original numbers. Here are some examples of the number of significant digits in a series of numbers.