FUNDAMENTAL STATISTICS FOR BEHAVIORAL SCIENCES

Fourth Edition/Robert B. McCall

FUNDAMENTAL STATISTICS FOR BEHAVIORAL SCIENCES

Fourth Edition

Robert B. McCall
Father Flanagan's Boys Town

Under the General Editorship of
Jerome Kagan
Harvard University

HBJ

Harcourt Brace Jovanovich, Publishers
San Diego New York Chicago Atlanta Washington, D.C.
London Sydney Toronto

ISBN: 0-15-529423-7

Library of Congress Catalog Card Number: 85-80085

Printed in the United States of America

COPYRIGHTS AND ACKNOWLEDGMENTS

AMERICAN CYANAMID COMPANY Excerpts from Wilcoxon, F., Katts, S., and Wilcox, R.A., *Critical Values and Probability Levels for the Wilcoxon Rank Sum Test and the Wilcoxon Signed Rank Test*. New York: American Cyanamid Company, 1965., and Wilcoxon, F., and Wilcox, R.A., *Some Rapid Approximate Statistical Procedures*, New York: Lederle Laboratories, 1964. All material reprinted by permission of the American Cyanamid Company.

BIOMETRIKA Table 2 of "Critical Values of the Coefficient of Rank Correlation for Testing the Hypothesis of Independence," *Biometrika*, December 1961, Vol. 48, Parts 3 and 4, pp. 444–48, by Gerald J. Glasser and Robert F. Winter. Reprinted by permission of Mrs. E.J. Snell, for and on behalf of the Biometrika Trustees.

ALLEN EDWARDS Table D, "Transformation of r to z_r from *Statistical Methods for the Behavioral Sciences* © 1954, published by Holt, Rinehart and Winston. Reprinted from Allen L. Edwards, *Statistical Methods*, Third Edition, 1973, by permission of the copyright owner, Allen L. Edwards.

IOWA STATE UNIVERSITY PRESS Table E, "Critical Values of F" from *Statistical Methods* by George B. Snedecor and William G. Cochran © 1967, published by the Iowa State University Press.

LONGMAN GROUP Tables III, VII, part of table IV, and XXXIII and IV are taken from Tables B, C, G, and K of Fisher and Yates: Statistical Tables for Biological, Agricultural and Medical Research published by Longman Group Ltd., London (previously published by Oliver and Boyd Ltd., Edinburgh) and by permission of the authors and publishers.

RANDOM HOUSE Table A, "Proportions of Area Under the Standard Normal Curve" and Table H, "Critical Values of the Mann-Whitney U for a Directional Test at .025 or a Nondirectional Test at .05" from *Fundamentals of Behavioral Statistics*, Third Edition by Richard Runyon and Audrey Haber. Copyright © 1976 by Newbery Award Records, Inc. Reprinted by permission of Random House, Inc.

WRIGHT PATTERSON AIR FORCE BASE Table F, "Distribution of the Studentized Range Statistic" as it appeared in *Statistical Principles in Experimental Design* published by McGraw-Hill in 1971. Reprinted by permission of Donald S. Clemm.

Formulas
Commonly Used in Computations

	FORMULA
Characteristics of Distributions of Scores	
Mean	$\overline{X} = \dfrac{\Sigma X_i}{N}$
Median	$M_d = L + \left[\dfrac{N/2 - n_b}{n_w}\right] i$
Mode	$M_o =$ most frequent X_i
Variance	$s^2 = \dfrac{\Sigma(X_i - \overline{X})^2}{N - 1} = \dfrac{N\Sigma X_i^2 - (\Sigma X_i)^2}{N(N-1)}$
Standard deviation	$s = \sqrt{s^2}$
Measures of Relative Standing	
Percentile point	$X = L + \left[\dfrac{P(N) - n_b}{n_w}\right] i$
Percentile rank	$P = \dfrac{n_w(X - L) + in_b}{Ni}$
Standard score	$z = \dfrac{X_i - \overline{X}}{s_x}$

Regression and Correlation

Given: $(\mathrm{I}) = N\Sigma X^2 - (\Sigma X)^2$ $\quad (\mathrm{II}) = N\Sigma Y^2 - (\Sigma Y)^2$ $\quad (\mathrm{III}) = N\Sigma XY - (\Sigma X)(\Sigma Y)$

Regression constants	
Slope	$b = \dfrac{(\mathrm{III})}{(\mathrm{I})}$
Intercept	$a = \overline{Y} - b\overline{X}$
Regression line	$\tilde{Y} = bX + a$
Standard error of estimate	$s_{y \cdot x} = \sqrt{\left[\dfrac{1}{N(N-2)}\right]\left[(\mathrm{II}) - \dfrac{(\mathrm{III})^2}{(\mathrm{I})}\right]}$

Correlation coefficient

$$r = \frac{(\mathrm{III})}{\sqrt{(\mathrm{I})(\mathrm{II})}}$$

Probability

Probability of A

$$P(A) = \frac{\#(A)}{\#(S)}$$

Probability of B given A (conditional probability)

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

Probability of A and B (intersection)

$$P(A \cap B) = P(A)P(B|A)$$

Probability of A or B (union)

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Permutations

$$_nP_r = \frac{n!}{(n-r)!}$$

Combinations

$$_nC_r = \frac{n!}{(n-r)!r!}$$

Binomial probability

$$P(r, n; p) = \frac{n!}{(n-r)!r!}p^r q^{n-r}$$

Elementary Parametric Techniques

Standard score forms

$$z_{\text{obs}} = \frac{X - \mu}{\sigma} \quad \text{or} \quad z_{\text{obs}} = \frac{\overline{X} - \mu}{\sigma_{\overline{x}}} = \frac{\overline{X} - \mu}{\sigma_x/\sqrt{N}}$$

$$t_{\text{obs}} = \frac{\overline{X} - \mu}{s_{\overline{x}}} = \frac{\overline{X} - \mu}{s_x/\sqrt{N}}$$

Confidence intervals for the mean

$$\text{Upper limit} = \overline{X} + t_{\alpha/2}(s_{\overline{x}})$$

$$\text{Lower limit} = \overline{X} - t_{\alpha/2}(s_{\overline{x}})$$

Difference between two independent groups

$$t_{\text{obs}} = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\left[\frac{(N_1 - 1)s_1^2 + (N_2 - 1)s_2^2}{N_1 + N_2 - 2}\right] \cdot \left[\frac{1}{N_1} + \frac{1}{N_2}\right]}}$$

$$df = N_1 + N_2 - 2$$

Difference between two correlated groups

$$t_{\text{obs}} = \frac{\Sigma D_i}{\sqrt{\frac{N\Sigma D_i^2 - (\Sigma D_i)^2}{N - 1}}} \qquad df = N - 1$$

(continued on back endpaper)

Preface

A veteran scholar of statistics once remarked that there are two ways to teach statistics—accurately or understandably. Disproving that "either-or" statement is the challenge that has guided my writing of this book.

As a textbook for the first course in applied statistics, *Fundamental Statistics for Behavioral Sciences* is used primarily by students majoring in psychology, the other social sciences, and education. In writing for this audience, my earliest and most basic decision was to emphasize the purpose, rationale, and application of important statistical concepts over rote memorization and the mechanical application of formulas. I believe that students at the introductory level, whether or not they plan to take advanced courses in statistics, are served better by a book that fosters an understanding of statistical logic than by one that stresses mechanics.

When understanding is emphasized, elementary statistics is neither dull nor mathematically difficult. *Fundamental Statistics for Behavioral Sciences* does not require much background in mathematics. The student need be familiar only with the thinking patterns learned in high school algebra and geometry; all relevant terms and operations are reviewed in Appendix 1. To be sure, the book contains many computations and problems to solve, but most statistical formulas rely heavily on simple arithmetic—addition, subtraction, multiplication, division, and the taking of square roots—and can be worked out quickly with the aid of a hand calculator. In addition, I have kept the data for computational problems simple so that the emphasis remains on the rationale and outcome of techniques instead of on calculation for its own sake.

Understanding a statistical concept or strategy can be a formidable mental exercise, but once it is done, details that torment mechanical learners become obvious deductions. The goal of understanding concepts is not hard to reach if students also understand as much of the mathematical reasoning as is within their grasp. Whenever possible, I have explained in mathematically simple terms the logic that undergirds the basic concepts and techniques, although a few items require advanced mathematics and must therefore be taken on faith at this level. Beginning in Chapter 3, optional tables that show full algebraic derivations and proofs supplement the test explanations. (These tables can be omitted without loss of continuity.)

For the sake of students, and contrary to the traditional practice of mathematical writers, I have included and explained every step in proofs, however "obvious." I have also avoided excessive use of symbols, since symbols require an extra mental translation and thus often confuse students. Deviation scores ($X_i - \overline{X}$), for instance, are *not* abbreviated by x. Further, each new symbol is carefully introduced and is frequently accompanied by its verbal equivalent.

Anyone who has analyzed his or her own data knows the anticipation that accompanies the final calculation of the r, t, or F hidden in a mass of numbers that took months to gather. Students, too—even though it is not their own data they are analyzing—can experience the excitement of seeing meaning emerge as they manipulate an apparently patternless collection of numbers. Yet they sometimes fail to see the fascination of statistical analysis because it is presented more as a numbers game played in a vacuum than as a crucial part of the scientific investigation of real phenomena.

For this reason, many end-of-chapter exercises and examples in the text are drawn from actual studies (modified for numerical simplicity). For example, the distinction between a correlation and a difference between means is demonstrated through findings that related the IQs of adopted children to those of the biological and adoptive mothers, and Freedman's work on the feeding behavior of dogs reared under indulged and disciplined conditions is used to illustrate interaction effects in the two-factor analysis of variance. Although most of these studies were performed by psychologists, many of them concern developmental and educational issues of interest to future teachers and school administrators.

As another means of showing that numbers can have real applications, I have tried to give students a feel for the behavior of a statistic by providing several data sets that display obvious contrasts. Before calculating an analysis of variance in Chapter 11, for example, I present a set of random numbers, introduce a main effect, and then add an interaction treatment effect. Many exercises ask students to alter a given set of scores in some way and to observe how the change affects the value of a statistic.

This Fourth Edition includes a number of changes, most of them the result of suggestions by users of the previous editions. The most obvious alteration is the new Chapter 14, on research design, which represents an expansion of the prologue in the previous edition. This information has been placed at the end of the text because many instructors feel the material is too difficult for students to grasp at the beginning of the course. Other additions include chapter summaries, a glossary of terms to complement the glossary of symbols, and short sections on how statistical results are presented in professional journals. The previous sections on individual comparisons, estimation, and confidence limits have been expanded—although they may be skipped by those who find them too elaborate for this level. Finally, the previous material on the second regression line, the relationship between F and t, and the table of squares and square roots has been eliminated.

In the Study Guide, the main section of each chapter is a semiprogrammed unit that reviews the basic terms, concepts, and computational routines described in the corresponding textbook chapter. The presentation is stripped of excess detail, and its tone is very concrete and applied. One feature that students have found particularly valuable has been retained: guided computational examples—step-by-step outlines that show how to organize the operations required in complicated calculations and that clarify the logical and computational details.

Individual instructors emphasize different aspects of elementary statistics, and many must select a subset of chapters that can be covered in the available time. The organization of the text is straightforward. Part 1 (Chapters 1–6) presents descriptive statistics, including central tendency, variability, relative position, regression, and correlation. Part 2 (Chapter 7–13) deals with inferential statistics, including probability, sampling distributions, the logic of hypothesis testing, elementary parametric tests, one- and two-factor analysis of variance, and a selection of nonparametric techniques. For most courses, the core chapters are 1–4 and 8–10; to these instructors may add at their option material on experimental design (Chapter 14), regression and correlation (Chapters 5 and 6), probability theory (Chapter 7), the analysis of variance (Chapters 11 and 12), and nonparametric tests (Chapter 13).

In addition to the colleagues and friends who contributed their advice to earlier editions, I want to thank Dr. Mark Appelbaum of the Psychometric Laboratory of the Department of Psychology at the University of North Carolina, Chapel Hill. He was my chief technical consultant, helping to blend my personal style with the demands of formal theory. I am indebted to the many whose commitment to the course and to the book led them to communicate directly with me over the years; their comments, criticisms, and suggestions stimulated many of the improvements in this new edition.

The revision of the Fourth Edition in particular has been guided by the suggestions of Drs. David Birch, University of Illinois; Richard Lindley, California State University at Fullerton; Walena Morse, West Chester University; Kelly Shaver, College of William and Mary; and Zita Tyer, George Mason University. Further, the job of producing a revision was made considerably easier by the assistance of Dr. Gaylon Oswalt, who reformulated the exercises in the text and the Study Guide, and Kathryn Whitmore, who prepared the manuscript and helped with the proofs.

I am grateful to the Literary Executor of the late Sir Ronald A. Fisher, F.R.S., to Dr. Frank Yates, F.R.S., and to Longman Group Ltd., London, for permission to reprint Tables III, IV, VIII, and part of XXXIII from their book *Statistical Tables for Biological, Agricultural and Medical Research* (Sixth Edition, 1974).

Special thanks go to my wife, Rozanne, for her encouragement and for the absence of complaints while she was temporarily widowed for this cause.

Robert B. McCall

Contents

PART 2
INFERENTIAL STATISTICS

PART 1

Descriptive Statistics

The Study of Statistics

CHAPTER 1

MOST OF US ENCOUNTER statistics every day. Averages and percentages, for example, are found in the sports sections of newspapers, in the weather forecast on television, and even on your grade reports. But, as you will see in this chapter, statistics is much more than just averages and percentages. An elementary knowledge of statistics is useful to almost everyone, but a detailed knowledge is absolutely essential to those who must read or conduct scientific research. In the social and behavioral sciences, such as psychology, sociology, education, and economics, research involves making observations or measurements on people, things, or events. **Statistics** is a set of procedures for describing those measurements and for making inferences about what is generally true. This chapter explains why a study of statistics is important, discusses the purpose of statistics and the kinds of measurements that can be made, and introduces a few elementary procedures for working with measurements.

WHY STUDY STATISTICS?

"Alcoholism is Genetic, Professor Claims," read the newspaper headline.[1] But exactly what does this mean—that whether or not a person becomes an alcoholic depends totally on his or her having a gene for alcoholism?

No, probably not. Other factors must also be taken into account. For example, the article goes on to say that children who are reared by at least one biological parent who is an alcoholic are almost six times more likely to become alcoholics than if no parent is an alcoholic.

But does that mean that if either the biological or rearing parent is an alcoholic, then the child will also become an alcoholic? No, it simply means that the likelihood is greater that the child will become an alcoholic. But just how great is the likelihood? Obviously, we need more information.

The article provides it; specifically, that there is an 8 to 14% likelihood that a child will become an alcoholic if no parent is an alcoholic but a 46 to 50% likelihood if one parent is an alcoholic. So even if a child has a parent who is an alcoholic, the chances are approximately one in two that he or she will not be one too.

What does this article have to do with statistics? It *is* statistics, at least partly. Statistics involves percentages, averages, "likelihoods," which are usually called probabilities, and much more. You encounter such statistics every day in newspaper articles like the one described above, in sports information, in the weather report, in economic reports and news summaries, in stock and bond information, and so forth. And of course science—especially the social and behavioral sciences—depends very heavily on statistics.

Beyond averages and percentages While you are familiar with averages and percentages, there are many other types of statistics as well, and sometimes you encounter them in the newspaper, not just in scientific journals. For example, you may read that the *median* family income in the United States in 1981 was $20,243. What is meant by the median? You may also read that the *average* family income in the same year was $22,787. How is the median different from the average, and what meaning does each statistic convey? You may also read that the average score on an IQ test is 100 and that the standard deviation is 16. What is the standard deviation, and how is it

[1] Inspired by, but not identical to, research in Denmark reported by M. A. Schuckit, D. A. Goodwin, and G. Winokur, "A Study of Alcoholism in Half Siblings," *American Journal of Psychiatry* 128 (1972): 1132–35.

useful in interpreting the mental ability of a child who scores 124, for example? And finally, you may read that children's IQ scores are more closely correlated with those of their biological parents than with those of their rearing parents, but that the IQs of the children are closer in value to the IQs of their rearing than of their biological parents. How is it possible for both these facts to be true? A knowledge of elementary statistics will help you answer these questions.

But, of course, a detailed knowledge of statistics is essential if you are going to study a social or behavioral science; if you must read scientific literature; or if you will ever conduct a scientific study. And if you go on to graduate school in psychology, sociology, education, economics, or some other social or behavioral discipline, you will probably be required to take more advanced courses in statistics.

Variability Social and behavioral scientists rely heavily on statistics because almost all numerical information, or data, in the social sciences contain variability. **Variability** refers to the fact that the scores or measurement values obtained in a study differ from one another, even when all the subjects in the study are assessed under the same circumstances. If your teacher were to give a test to all the students in your statistics class, you would not all obtain the same score. Similarly, all rats tested in a multiple-T maze do not make the same number of errors. This dissimilarity in score values is variability.

There is variability in social science data for at least three reasons. First, the units (people, usually) that social scientists study are rarely identical to one another. A chemist or a physicist assumes that each of the units under study (molecules, atoms, electrons, and so forth) is identical in its composition and behavior to every other unit of its kind. But the social scientist cannot assume that all people will respond identically in a given situation. In fact, the social scientist can count on the fact that they will *not* respond in the same way. The first major source of variability, then, is **individual differences** in the behavior of different subjects.

Second, social scientists cannot always measure as accurately as they would like the attribute or behavior they wish to study. Scores on a classroom examination are supposed to be a measure of how much students have learned, but even a good test has flaws that make it an imperfect measure of learning. Perhaps some questions are ambiguous, or perhaps one part of the material is overemphasized. These and all other factors besides knowledge of the material that influence test performance constitute **measurement error** and contribute to variability. Almost all behavioral measurements contain some measurement error.

Third, even a single unit (person, animal) usually will not respond exactly the same way on two different occasions. If you have a person step on a precise scale 10 times in succession, the scale will probably record the same weight each time. But if you give a person several opportunities to rate the attractiveness of advertising displays or the degree of aggressiveness in filmed episodes of each of 15 nursery-school children, he or she will most likely not assign the same scores on each occasion. This source of variability is called **unreliability**.

Variability means that less faith can be placed in a single score than one would like. Therefore, it may be necessary either to measure a single subject several times or to assess many different subjects or events in order to obtain data that faithfully reflect the characteristics being studied.

Individual differences, measurement error, and unreliability are the major reasons that it is often difficult to draw accurate conclusions from behavioral and social science data. It is the task of statistics to quantify the variability in a set of measurements; to describe the data for a group of subjects, despite the variability inherent in the measurements; and to derive precisely stated and consistent decisions about the results by quantifying the uncertainty produced by variability.

DESCRIPTIVE AND INFERENTIAL STATISTICS

Statistics is the study of methods for describing and interpreting this quantitative information, including techniques for organizing and summarizing data and techniques for making generalizations and inferences from data. The first of these two broad classes of methods is called descriptive statistics, and the second is called inferential statistics.

Descriptive statistics refers to procedures for organizing, summarizing, and describing quantitative information or data. Most people are familiar with some statistics of this sort. The baseball fan is accustomed to checking over a favorite player's batting average; the sales manager relies on charts showing the sales distribution and cost efficiency of an enterprise; the head of a household may consult tables showing the average domestic expenditures of families of comparable size and income; and the actuary possesses charts outlining the life expectancies of people in various professions. These are relatively simple statistical tools for characterizing data, but additional techniques are available to describe such things as the extent to which measured values deviate from one another and the relationship between individual differences in performance of two different kinds. For example, one can describe statistically the degree of relationship between the scores of a group of students on a college entrance examination and their later performance in college.

The second major class of statistics, **inferential statistics**, includes methods for making inferences about a larger group of individuals on the basis of data concerning a small group. For example, people who develop diabetes in adulthood are often treated with the drug insulin. The insulin is thought not only to regulate blood sugar level, preventing coma and possibly death, but to retard the progressive damage to blood vessels that is the major cause of blindness, heart attacks, disability, and death among diabetics. Suppose a test had been made to determine whether people who were treated with insulin suffered fewer of these major consequences than patients who were treated with a special diet thought to be beneficial, but not with insulin.[2] Two groups, each consisting of 200 diabetic volunteers judged medically able to participate in the study, followed a special diet and received "medication." For one group the medication was insulin, but for the other group it was a substance that was known to have no effect (a placebo). Both groups of patients were observed for 9 to 11 years. Suppose that during those years 56 of the insulin group and 63 of the non-insulin group suffered at least one of the known consequences of diabetes. The question is: Did the insulin treatment and diet together do more to help prevent undesirable complications than the special diet alone?

Note that whether the 200 people who received insulin had fewer problems than the 200 who took the placebo is *not* the question. The results of the experiment itself have answered that—56 versus 63 patients developed later problems. What we would really like to know is whether insulin is likely to have this effect generally for all such people, even though we must make that decision on the basis of results determined with only 400 people. We must make a decision about a larger group based upon observations of a much smaller group. That is, we want to draw a statistical inference.

However, we feel some uncertainty about deciding whether insulin is generally helpful or not. While 56 cases are certainly fewer than 63, perhaps such a difference is the result of chance. Imagine for a moment that the insulin in fact had no effect whatsoever. Even in this case, we would not expect *exactly* the same number of complications in each group, because one group of 200 people and the circumstances of their disease are not identical in every way to another group of 200. One group might have 58 cases of complications and the other 61, or perhaps 56 and 63. So the question we wish to answer through the use of inferential statistics is: How much of a difference between these two groups must there be before it is safe to decide that insulin helps?

The uncertainty that we feel can be quantified, and this quantification relies on **probability**. When we say that the likelihood of a fair coin coming up heads is .50, we have made a probability statement that quantifies our uncertainty about the outcome of flipping the coin. The probability value really says that in an uncountable number of flips, .50 (or 50%) will turn up heads. It is in this sense that probability quantifies our uncertainty about what will happen in a large number of cases.

Probability is used in statistical inference in the same way as in daily life. For instance, you might consider the probability of rain in deciding whether or

[2] Inspired by, but not identical to, the study reported by Genell L. Knatterud, Christian R. Klimt, Marvin E. Levin, Maynard E. Jacobson, and Martin G. Goldner, "Effects of Hypoglycemic Agents on Vascular Complications in Patients with Adult-Onset Diabetes," *Journal of the American Medical Association* 240 (1978): 37–42.