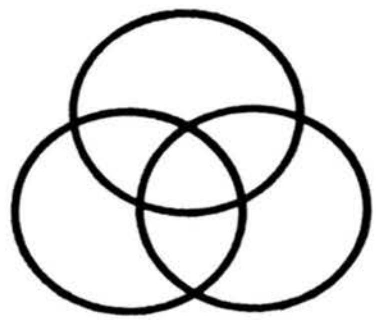
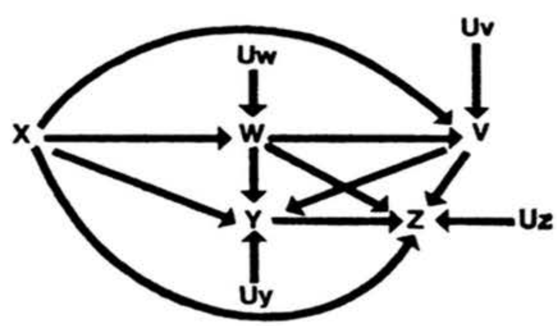
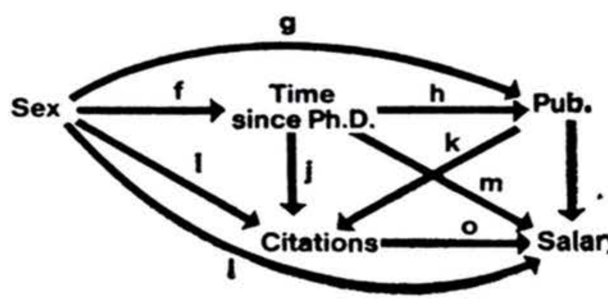
Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences

Second Edition







Jacob Cohen
Patricia Cohen

Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences

Second Edition

Page Intentionally Left Blank

Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences

Second Edition

Jacob Cohen
New York University

Patricia Cohen
New York State Psychiatric Institute
and
Columbia University School of Public Health

Psychology Press
Taylor & Francis Group
New York London

First Published by
Lawrence Erlbaum Associates, Inc., Publishers
10 Industrial Avenue
Mahwah, New Jersey 07430

Transferred to Digital Printing 2009 by Psychology Press
270 Madison Ave, New York NY 10016
27 Church Road, Hove, East Sussex, BN3 2FA

Library of Congress Cataloging in Publication Data

Cohen, Jacob, 1923–
Applied multiple regression/correlation analysis for the behavioral sciences.

Bibliography: p.
Includes index.
1. Regression analysis. 2. Correlation (Statistics)
I. Cohen, Patricia. II. Title. III. Title: Multiple regression/correlation analysis for the behavioral sciences. [DNLM: 1. Behavioral sciences 2. Regression analysis. HA 21.3.C678a]
HA31.3.C63 1983 519.5'36 83-11487
ISBN 0-89859-268-2

Publisher's Note
The publisher has gone to great lengths to ensure the quality of this reprint but points out that some imperfections in the original may be apparent.

to **Gideon Moses Cohen**

(another collaborative product,
but in no need of revision)

Page Intentionally Left Blank

Contents

PART II: THE REPRESENTATION OF INFORMATION IN INDEPENDENT VARIABLES

PART III: APPLICATIONS

APPENDICES

Preface to the Second Edition

The seven years since the publication of the first edition have been fat ones for multiple regression/correlation as a general data-analytic system ("new-look" MRC, in short). The behavioral and social science journals have carried hundreds of pages on methodological issues in MRC, (much of it on how–when–whether it can replace other methods), and, increasingly, research reports that employ it. Several MRC textbooks and chapter-length treatments have appeared, and MRC interest groups have formed as part of specialized scientific organizations. "New-look" MRC has been applied across a spectrum that reaches from the molecular end of psychology through the molar end of sociology, and has been particularly popular in education, the evaluation of intervention programs, drug research, and psychiatric epidemiology. Its obvious relevance to "meta-analysis" has not been overlooked.

While much of the "nuts and bolts" in the original edition remains intact, there has been a fundamental change in outlook. The relevance of MRC to the study of causality, dimly hinted at here and there in the first edition, now emerges as the central principle. From the very beginning in the presentation of the two-variable regression equation, through the interpretation of patterns of association with two independent variables, alternative analytic strategies, and setwise hierarchical analysis, the intimate relationship of MRC to the formal analysis of causal models is described and illustrated. After the methods of representing information as data for MRC are presented, a chapter is devoted to the analysis of causal models, and simple working methods employing systems of regression equations are provided.

The detailed exposition of the analysis of covariance is preceded by a causal models analysis of the types of research design that employ this method. Throughout, the exposition reflects our conviction that the valid analysis of nonexperimental data can proceed only when it is in keeping with the principles and insights of the analysis of causal models.

Another change we believe to be important does not occur in the body of the text, but appears as Appendix 4. It is a reprinting, with minor revisions, of a newly derived multivariate generalization of MRC, one that handles sets and partialled sets of variables as *dependent* variables. Although we have found this method most useful in our work and believe that it is the latest word in quantitative methodology, its novelty and our academic caution (as well as our unfailing modesty) dictate that it not be part of the central expository framework (although it has led us to modify the material on canonical analysis, which it seeks to largely replace).

Other new material includes a test of the omnibus null hypothesis that all of the correlations in a matrix are zero in the population, the analysis of residuals, the analysis of conditional missing data (such as result from such survey items as "If not presently married, skip to item 16"), and a detailed analysis of the role of scaling in the interpretation of interaction regression coefficients.

The goals, style, tone, and emphasis remain the same. We continue to intend the book to serve both as a textbook for students and a handbook for researchers. The nonmathematical, intuitive, applied, data-analytic features remain. The emphasis on the use of MRC in the service of scientific explanation rather than that of forecasting continues. While the latter is not neglected, its limitations are noted in a new section on unit weighting and other alternatives to regression weights in prediction. As before, the exposition is heavily laced with worked examples; one, on factors associated with academic salaries, is carried through the book and culminates in its use to exemplify the complete analysis of a causal model.

As always, we first acknowledge the many students, colleagues, and researchers whose response to the first edition and other interaction with us taught us at least as much as we taught them. Again, our colleagues at meetings of the Society of Multivariate Experimental Psychology served as a constructively critical forum for some of the new material and our general approach. Among them, John Loehlin provided helpful comments on Chapters 3 and 9 as did the anonymous referees of our journal, Multivariate Behavioral Research, on the set correlation paper (Appendix 4). We thank Gregory Muhlin for help with the computer program information in Appendix 3, and E. L. Struening for his general support. Larry Erlbaum has earned a special debt of gratitude for making it all so easy. Annette Friedner and Detra Allen did most of the typing, and beautifully.

JACOB COHEN
PATRICIA COHEN

Preface to the First Edition

This book had its origin about a dozen years ago, when it began to become apparent to the senior author that there were relationships between regression and correlation on the one hand and the analysis of variance on the other which were undreamed of (or at least did not appear) in the standard textbooks with which he was familiar. On the contrary, the texts, then as now, treated these as wholly distinct systems of data analysis intended for types of research which differed fundamentally in design, goals, and types of variables. Some research, both statistical and bibliographic, confirmed the relationships noted, and revealed yet others. These relationships served to enrich both systems in many ways, but it also became clear that multiple regression/correlation was potentially a very general system for analyzing data in the behavioral sciences, one that could incorporate the analysis of variance and covariance as special cases. An article outlining these possibilities was published in the *Psychological Bulletin* (Cohen, 1968), and the volume and sources of reprint requests and several reprintings suggested that a responsive chord had been struck among behavioral scientists in diverse areas. It was also obvious that for adequacy of both systematic coverage and expository detail, this book-length treatment was needed.

In 1969, the authors were married and began a happy collaboration, one of whose chief products is this book. (Another is saluted on the dedication page.) During the preparation of the book, the ideas of the 1968 paper were expanded, further systematized, tried out on data, and hardened in the crucible of our teaching and consulting. We find the system which has evolved surprisingly easy to teach and learn, and this book is an effort to so embody it. We omit from this preface, except incidentally, a consideration of this book's scope, orientation, and organization, since Chapter 1 is largely devoted to these issues.

To describe the primary audience for whom this book is intended requires two dimensions. Substantively, this book is addressed to behavioral and social scientists. These terms have no sharply defined reference, but we intend them in the most inclusive sense to include the academic sciences of psychology, sociology, economics, branches of biology, political science, and anthropology, and also various applied research fields: education, clinical psychology and psychiatry, epidemiology, industrial psychology, business administration, social work, and political/social survey, market and consumer research. Although the methods described in this book are applicable in other fields (for example, industrial engineering, agronomy), our examples and atmospherics come from behavioral–social science.

The other dimension of our intended audience, amount of background in statistics and research, covers an equally broad span. This book is intended to be both a textbook for students and a manual for research workers, and thus requires a somewhat different approach by these two readerships. However, one feature of this book will be appreciated by a large majority of both groups of readers: its orientation is nonmathematical, applied, and "data-analytic." This orientation is discussed and justified in the introductory chapter (Section 1.2.1) and will not be belabored here. Our experience has been that with few exceptions, both students and research practitioners in the behavioral and social sciences approach statistics with considerable wariness (to say the least), and require a verbal–intuitive exposition, rich in redundancy and concrete examples. This we have sought to supply.

As a textbook, whether used in a course at the graduate or advanced undergraduate level, it is assumed that the students have already had a semester's introductory statistics course. Although Chapter 2 begins "from scratch" with bivariate correlation and regression, and reviews elementary statistical concepts and terminology, it is not really intended to be a thorough, basic exposition, but rather to refresh the reader's memory. Students without a nodding acquaintance with the analysis of variance will find some portions of Chapter 1 difficult; returning to this material later in the course should clear matters up. Increasingly, statistical offerings in graduate programs are so organized as to include a course in correlation/regression methods. This book is intended to serve as a text for such courses. It may also be used in courses in multivariate methods—although largely devoted to multiple regression/correlation analysis, the final chapter links it to and reviews the other multivariate methods.

As a manual, this book provides an integrated conceptual system and practical working methods for the research worker. The last five years has seen a rapidly growing interest in multiple regression/correlation methods, reflected in journal articles and books addressed to psychologists and sociologists. Much of this material is valuable, while some of it is misguided or simply incorrect. Some of the more valuable contributions are presented mathematically, thus limiting their access. Taken as a whole, the recent literature is lacking in the combination of an integrated conceptual system with easily understood practical working methods which is necessary for the method to realize its potential as a general data-analytic system. We have tried to provide this. Chapter 1 begins with an outline of this system, and was written primarily with the experienced research worker or advanced graduate student in mind. He or she will find much of Chapters 2 and 3 elementary, but they are worth skimming, since some of the topics are treated from a fresh perspective which may be found insight provoking. Chapter 4 considers sets of independent variables as units of analysis, and is basic for much of what follows. Beyond that, he may follow his specific interests in the chapters and appendices by reference to the table of contents and a carefully prepared index. To the stat buff or teacher, we recommend reading the chapters in order, and the appendices at the point in the text where they are referenced.

We acknowledge, first of all, the many students, colleagues and researchers seeking counsel whose stimulation so importantly shaped this book. A small subset of these are Elmer L. Struening, Mendl Hoffman, Joan Welkowitz, Claudia Riche, and Harry Reiss, but many more could be named. Special thanks are due to the members of the Society of Multivariate Experimental Psychology for the useful feedback they supplied when portions of the book were presented at their annual meetings during the last few years. We are very grateful to Joseph L. Fleiss for a painstaking technical critique from which the book greatly profited. Since we remained in disagreement on some points, whatever faults remain are our sole responsibility. Gerhard Raabe provided valuable advice with regard to the material on computer programs in Appendix 3. Patra Lindstrom did a most competent job in typing the manuscript.

JACOB COHEN
PATRICIA COHEN

Page Intentionally Left Blank

Applied Multiple
Regression/Correlation Analysis
for the Behavioral Sciences

Second Edition

Page Intentionally Left Blank

PART I

BASICS

Page Intentionally Left Blank

1

Introduction

1.1 MULTIPLE REGRESSION/CORRELATION AS A GENERAL DATA-ANALYTIC SYSTEM

1.1.1 Overview

Multiple regression/correlation analysis (MRC) is a highly general and therefore very flexible data-analytic system that may be used whenever a quantitative variable (the dependent variable) is to be studied as a function of, or in relationship to, any factors of interest (expressed as independent variables). The sweep of this statement is quite intentional:

1. The form of the relationship is not constrained; it may be simple or complex, for example, straight line or curvilinear, general or conditional, or combinations of these possibilities.
2. The nature of the research factors expressed as independent variables is also not constrained; they may be quantitative or qualitative, main effects or interactions in the analysis of variance (AV) sense, or covariates as in the analysis of covariance (ACV). They may be characterized by missing data. They may be correlated with each other, or uncorrelated (as in balanced factorial design AV). They may be naturally occurring ("organismic") properties like sex or diagnosis or IQ, or they may be consequences of planned experimental manipulation ("treatments"). They may be single variables or groups of variables. In short, virtually any information whose bearing on the dependent variable is of interest may be expressed as research factors.

The MRC system presented in this book has other properties that make of it a powerful analytic tool: it yields measures of the magnitude of the "whole" relationship of a factor to the dependent variable, as well as of its partial (unique, net) relationship, that is, its relationship over and above that of other research factors (proportions of variance and coefficients of correlation and regression). It also comes fully equipped with the necessary apparatus for statistical hypothesis testing, estimation, and power analysis. Last, but certainly not least, it is a major tool in the methods of causal analysis.

In short, and at the risk of sounding like a television commercial, it is a versatile, all-purpose system for analyzing the data of the behavioral, social, and biological sciences and technologies.

We are, of course, describing the "new-look" MRC that has developed over the past 2 decades, not the stereotyped traditional application that was largely limited to the psychotechnological task of forecasting outcomes in educational or personnel selection and vocational guidance. The very terminology betrays this preoccupation: *Criterion* (dependent) variables are *predicted* by "predictor" (independent) variables, correlations between *predictors* and *criterion* are called *validity coefficients,* and accuracy is assessed by means of the "standard error of prediction." Even when put to other uses, stereotypy is induced either implicitly or explicitly by limiting MRC to straight-line relationships among equal-interval scales on which the observations are assumed to be normally distributed. In this narrow view, MRC takes its place as one of a group of specialized statistical tools, its use limited to those occasional circumstances when its unique function is required and its working conditions are met.

Viewed from this traditional perspective, it is hard to see why anyone would want a whole textbook devoted to MRC, a monograph for specialists perhaps, but why a textbook? No such question arises with regard to textbooks entirely devoted to the analysis of variance and covariance, because of its presumed generality. This is ironic, since, as we will show, AV/ACV is in fact a special case of MRC!

Technically, AV/ACV and conventional multiple regression analysis are special cases of the "general linear model" in mathematical statistics.[1] The MRC system of this book generalizes conventional multiple regression analysis to the point where it is essentially equivalent to the general linear model. It thus follows that any data analyzable by AV/ACV may be analyzed by MRC, while the reverse is not the case. This is illustrated, for example, by the fact that when one seeks in an AV/ACV framework to analyze a factorial design with unequal cell frequencies, the nonindependence of the factors necessitates moving up to the more general multiple regression analysis to achieve an exact solution.

Historically, MRC arose in the biological and behavioral sciences around the turn of the century in the study of the natural covariation of observed characteristics of samples of subjects (Galton, Pearson, Yule). Somewhat later, AV/ACV grew out of the analysis of agronomic data produced by controlled variation of treatment conditions in manipulative experiments (Fisher). The systems developed in parallel, and from the perspective of the research workers who used

[1]For the technically minded, we point out that it is the "fixed" version of these models to which we address ourselves, which is the way they are most often used.

them, largely independently of each other. Indeed, MRC, because of its association with nonexperimental, observational, survey-type research, came to be looked upon as less scientifically respectable than AV/ACV, which was associated with experiments. The recent development of causal analysis, formal systems of inference based on nonexperimental data, with its heavy dependence on regression analysis, has tended to offset this onus.

Close examination suggests that this guilt (or virtue) by association is unwarranted—the result of the confusion of data-analytic method with the logical considerations which govern the inference of causality. Experiments in which different treatments are applied to randomly assigned groups of subjects permit unambiguous inference of causality, while the observation of associations among variables in a group of randomly selected subjects does not. Thus, the finding of significantly more cell pathology in the lungs of rats reared in cigarette smoke-filled environments than for normally reared control animals is in a logically superior position to draw the causal inference than is the finding that, for a random sample of postmortem cases, the lung cell pathology is significantly higher for divorced men than for married men. But each of these researches may be analyzed by either AV (a simple pathology mean difference and its t test) *or* MRC (a simple correlation between group membership and pathology and its identical t test). The logical status of causal inference is a function of how the data were produced, not how they are analyzed. Yet it is not surprising that the greater versatility of MRC has made it the vehicle of formal causal analysis.

Authors who make sharp theoretical distinctions between correlation and fixed-model AV (or fixed-model regression) are prone to claim that correlation and proportion of variance (squared correlation) measures lack meaning for the latter because these measures depend on the specific levels of the research factor chosen (fixed) by the investigator and the (fixed) number of cases at each level. Concretely, consider an experiment where random samples of subjects are exposed to several levels of magnitude of a sensory stimulus, each sample to a different level, and their responses recorded. Assume that, over the purists' objections, we compute a correlation between stimulus condition and response, and find it to be .70, that is, about half ($.70^2 = .49$) of the response variance is accounted for by the stimulus conditions. They would argue, quite correctly, that the selection of a different set of stimulus values (more or less varying, occurring elsewhere in the range), or a different distribution of relative sample sizes, would result in a larger or smaller proportion of the response variance being accounted for. Therefore, they would argue, the .49 (or .70) figure can not be taken as an estimate of "the relationship between stimulus and response" for this sensory modality and form of response. Again, we must agree. Therefore, they would finally argue, these measures are meaningless. Here, we beg to disagree. We find such measures to be quite useful, provided that their dependence on the levels and relative sample sizes of the research factor is understood. When necessary, one simply attaches to them, as a condition or qualification, the distribution of the research factor. We find such qualifications no more objectionable, in principle, than the potentially many others (apparatus, tests, time of day, subjects, experimenters) on which research results may depend. Such measures, qualified as necessary, may mean more or less, depending on substantive considerations, but they are hardly meaningless. (For an example where the research factor is religion, and further discussion of this issue, see Section 5.3.1)

On the contrary, one of the most attractive features of MRC is its automatic provision of regression coefficients, proportion of variance, and correlation measures of various kinds. These are measures of "effect size," of the magnitude of the phenomena being studied. We venture the assertion that, despite the preoccupation (some critics would substitute "obsession") of the behavioral and social sciences with quantitative methods, the level of consciousness in many areas of just how big things are is at a surprisingly low level. This is because concern about the statistical significance of effects (whether they exist at all) has tended to preempt attention to their magnitude. That significant effects may be small, and nonsignificant ones large, is a truism. Although not unrelated, the size and statistical significance of effects are logically independent features of data from samples. Yet many research reports, at least implicitly, confuse the issues of size and statistical significance, using the latter as if it meant the former. At least part of the reason for this is that traditional AV/ACV yields readily interpretable F and t ratios for significance testing, but offers differences between group means as measures of effect size.[2] Now, a difference between means is a reasonably informative measure of effect size when the dependent variable is bushels of wheat per acre, or dollars of annual income, or age at time of marriage. It is, however, less informative when the dependent variable is a psychological test score, a sociological index, or the number of trials to learn a maze. Many of the variables in the behavioral and social sciences are expressed in units that are arbitrary, ad hoc, or unfamiliar. To report, for example, that law students show a 9.2-point higher mean than medical students on a scale measuring attitude toward the United Nations is to convey very little about whether this constitutes a large or trivial difference. However, to report that the law student–medical student distinction accounts for 4% of the attitude score variance conveys much more. Further, to report that the law students' mean on another scale, attitude toward public service, is 6.4 points higher than the medical students' mean not only fails to convey a useful sense of the size of this difference (as before), but is not even informative as to whether this is a smaller or larger difference than the other, since the units of the two scales are not directly comparable. But reporting that this distinction accounts for 10% of the public service attitude score variance not only expresses the effect size usefully, but *is*

[2]This is not to say that the more useful measures of proportion of variance have not been proposed in AV contexts; see, for example, Hays (1981), Cohen (1965, 1977), and Section 5.3.4. But they are neither integral to the AV tradition nor routinely presented in research reports where the data are analyzed by AV/ACV.

comparable to the 4% found for the other variable. Since various types of proportion of variance, that is, the squares of simple, multiple, partial, and semipartial correlation coefficients, may be routinely determined in the MRC system, the latter has "built-in" effect size measures that are unit-free and easily understood and communicated. Each of these comes with its significance test value for the null hypothesis (F or t), and no confusion between the two issues of *whether* and *how much* need arise.

1.1.2 Multiple Regression/Correlation and the Complexity of Behavioral Science

The greatest virtue of the MRC system is its capacity to mirror, with high fidelity, the complexity of the relationships that characterize the behavioral sciences. The word *complexity* is itself used here in a complex sense to cover several issues.

Multiplicity of Influences

The behavioral sciences inherited from older branches of empirical inquiry the simple experimental paradigm: vary a single presumed causal factor (C) and its effects on the dependent variable (Y), while holding constant other potential factors. Thus, $Y = f(C)$; variation in Y is a function of controlled variation in C. This model has been, and continues to be, an effective tool of inquiry in the physical sciences and engineering, and in some areas of the behavioral sciences. Probably because of their much higher degree of evolution, functional areas within the physical sciences and engineering typically deal with a few distinct causal factors, each measured in a clear-cut way, and each in principle independent of others.

However, as one moves from the physical sciences through biology and across the broad spectrum of the behavioral sciences ranging from physiological psychology to cultural anthropology, the number of potential causal factors increases, their representation in measures becomes increasingly uncertain, and weak theories abound and compete. Consider a representative set of dependent variables: epinephrine secreted, distant word associations, verbal learning, school achievement, psychosis, anxiety, aggression, attitude toward busing, income, social mobility, birth rate, kinship system. A few moments' reflection about the causal nexus in which each of these is embedded suggests a multiplicity of factors, and possibly further multiplicity in how any given factor is represented. Given several research factors C, D, E, etc., to be studied, one might use the single-factor paradigm repeatedly in multiple researches, that is, $Y = f(C)$, then $Y = f(D)$, then $Y = f(E)$, etc. But MRC makes possible the use of paradigms of the form $Y = f(C, D, E, \text{etc.})$, which are far more efficient than the strategy of studying multiple factors one at a time. Moreover, causal analysis, utilizing interlocking regression equations as formal models, achieves an even greater degree of appositeness to complex theories.

Correlation among Research Factors and Partialling

A far more important type of complexity than the sheer multiplicity of research factors lies in the effect of relationships among them.

The simpler condition is that in which the factors $C, D, E, \ldots$ are statistically unrelated (orthogonal) to each other, as is the case in true experiments where they are under the experimenter's manipulative control. The overall importance of each factor (for example, the proportion of Y variance it accounts for) can be unambiguously determined, since its orthogonality with the other factors assures that its effects on Y can not overlap with the effects of the others. Thus, concretely, consider a simple experimental inquiry into the proposition "don't trust anyone over 30" in which the persuasibility (Y) of male college students is studied as a function of the apparent age (C: C_1 = under 30, C_2 = over 30) and sex (D: D_1 = male, D_2 = female) of the communicator of a persuasive message. The orthogonality of the research factors C and D is assured by having equal numbers of subjects in the four "cells" (C_1D_1, C_1D_2, C_2D_1, C_2D_2); no part of the difference in overall Y means for the two communicator ages can be attributed to their sexes, and conversely, since the effect of each factor is balanced out in the determination of the other. If it is found that C accounts for 10% of the Y variance, and D for 5%, no portion of either of these amounts can be due to the other factor. It thus follows that these amounts are additive: C and D together account for 15% of the Y variance.[3]

Complexity arises when one departs from manipulative experiments and the orthogonality of factors which they make possible. Many issues in behavioral sciences are simply inaccessible to true experiments, and can only be addressed by the systematic observation of phenomena as they occur in their natural flux. In nature, factors which impinge on Y are generally correlated with each other as well. Thus, if persuasibility (Y) is studied as a function of authoritarianism (C), intelligence (D), and socioeconomic status (E) by surveying a sample with regard to these characteristics, it will likely be found that C, D, and E are to some degree correlated with each other. If, taken singly, C accounts for 8%, D 12%, and E 6% of the Y variance, because of the correlations among these factors, it will not be the case that together, in an MRC analysis, they account for 8 + 12 + 6 = 26% of the Y variance. It will almost certainly be less (in this case, but may, in general, be more—see Section 3.4). This is the familiar phenomenon of redundancy among correlated explanatory variables with regard to what they explain. The Y variance accounted for by a factor is overlapped to some degree

[3]The reader familiar with AV will recognize this as a balanced 2 × 2 factorial design. To avoid possible confusion, it must be pointed out that the orthogonality of C and D is a fact which is wholly independent of the possibility of a $C \times D$ interaction. Interactions are research factors in their own right and in balanced designs are also orthogonal to all other research factors. If the $C \times D$ interaction were to be included as a third factor and found to account for 2% of the variance, this amount is wholly its own, and all three factors combined would account for 17% of the Y variance. See the section "General and Conditional Relationships," and Chapter 8, which is devoted entirely to interactions.

with others. This in turn implies the concept of the variance accounted for by a factor *uniquely*, relative to what is accounted for by the other factors. In the above example, these unique proportions of Y variance may turn out to be: C 4%, D 10%, and E 0%. This is a rather different picture than that provided by looking at each factor singly. For example, it might be argued that E's apparent influence on Y when appraised by itself is "spurious," being entirely attributable to its relationship to C and/or D. Detailed attention to the relationships *among* the causal variables and how these bear on Y is the hallmark of causal analysis, and may be accomplished by MRC.

MRC's capability for assessing unique variance, and the closely related measures of *partial* correlation and regression coefficients it provides, is perhaps its most important feature, particularly for observational (nonexperimental) studies. Even a small number of research factors define many alternative possible causal systems or theories. Selection among them is greatly facilitated by the ability, using MRC, of partialling from the effects of any research factor those of any desired set of other factors. It is a copybook maxim that no correlational method can establish causal relations, but certain causal alternatives may be invalidated by the skillful use of this feature of MRC. It can show whether a set of observational data for Y and the correlated research factors C and D are consistent with any of the following possibilities, which are also expressed, parenthetically, in the language of causal analysis:

1. C and D each bears causally on Y (each has a *direct* effect on Y.)
2. C is a surrogate for D in relationship with Y, that is, when D is partialled from C, the latter retains no variance in Y. (C has no direct effect on Y.)
3. D suppresses the effect of C on Y, that is, when D is partialled from C, the unique variance of C in Y is greater than the proportion it accounts for when D is ignored (see the discussion of "suppression" in Section 3.4). (D represents a cause that is correlated with C but acts on Y in a direction opposite from C.)

The possibility for complexity in causal structures is further increased when the number of research factors increases beyond two, yet the partialling inherent in MRC is a powerful adjunct to good theory (i.e., causal models) for disentangling them. Further, partialling is at the base of a series of data-analytic procedures of increasing generality, which are realizable through MRC: general ACV, the Analysis of Partial Variance (see Chapter 10). Most generally, it is the partialling mechanism more than any other feature which makes it possible for the MRC system to mirror the complexity of causal relationships encountered in the behavioral sciences.

Form of Information

The behavioral and social sciences utilize information in various forms. One form which research factors may take is quantitative, and of any of the following levels of measurement (Stevens, 1951):

1. *Ratio Scales.* These are equal interval scales with a true zero point, making such ratio statements as "J has twice as much X as K" sensible. X may here be, for example, inches, pounds, seconds, foot-candles, voltage, size of group, dollars, distance from hospital, years in prison, or literacy rate.

2. *Interval Scales.* These have equal intervals but are measured from an arbitrary zero point, that is, the value of X that denotes absence of the property is not defined. Most psychological measures and sociological indices are at this level, for example, the scores of tests of intelligence, special abilities, achievement, personality, temperament, vocational interest, and social attitude. A physical example is temperature measured in Fahrenheit or Centigrade units.

3. *Ordinal Scales.* Only the relative position within a collection are signified by the values of ordinal scales, neither conditions of equal intervals nor a true zero obtaining. Whether simple rank order values are used, or they are expressed as percentiles, deciles, or quartiles, these properties of ordinal scales are the same.

The above scheme is not exhaustive of quantitative scales, and others have been proposed. For example, psychological test scores are unlikely to measure with exactly equal intervals and it may be argued that they fall along a continuum between interval and ordinal scales. Also, some rating scales frequently used in applied psychological research are not covered by the Stevens scheme since they have a defined zero point but intervals of dubious equality, for example, 0-never, 1-seldom, 2-sometimes, 3-often, 4-always.

Nominal Scales. Traditional MRC analysis was generally restricted to quantitative scales with (more or less) equal intervals. But much information in the behavioral sciences is not quantitative at all, but qualitative or categorical, or, using Steven's (1951) formulation, measured on "nominal" scales. Whether they are to be considered a form of measurement at all is subject to debate, but they undoubtedly constitute information. Some examples are: ethnic group, place of birth, religion, experimental group, marital status, psychiatric diagnosis, type of family structure, choice of political candidate, sex. Each of these represents a set of mutually exclusive categories which accounts for all the cases. The categories of true nominal scales represent distinguishable qualities, without natural order or other quantitative properties. Thus, nominal scales are sets of groups which differ on some qualitative attribute.

When research factors expressed as nominal scales are to be related to a dependent variable Y, past practice has been to put them through the mill of AV, whose grist is Y values organized into groups. But the qualitative information which constitutes nominal scales may be expressed quantitatively, and used as independent variables in MRC. (Chapter 5 is devoted to this topic, and answers such questions as, "How do you score religion?")

The above does not exhaust the forms in which information is expressed, since mixtures of scale types and other irregularities occur in practice. For example, interviews and questionnaires often require for some items the provision of categories for "does not apply" and/or "no response"; some questions are asked only if a prior question has had some specified response. As uninformative as such categories may seem at first glance, they nevertheless contain information and are capable of expression in research factors (see Chapter 7).

The above has been presented as evidence for that aspect of the complexity of the behavioral sciences which resides in the great variety of forms in which their information comes. Beginning with Chapter 5, we shall show how information in any of these forms may be used as research factors in new-look MRC. The traditional restriction of MRC to equal interval scales will be shown to be quite unnecessary. The capacity of MRC to use information in almost any form, and to mix forms as necessary, is an important part of its adaptive flexibility. Were it finicky about the type of input information it could use, it could hardly function as a *general* data-analytic system.

Shape of Relationship

When we come to scrutinize a given relationship expressed by $Y = f(C)$, it may be well described by a straight line on the usual graph, for example, if Y and C are psychological tests of abilities. Or, adequate description may require that the line be curved, for example, if C is age, or number of children, such may be the case. Or, the shape may not be definable, as when C is a nominal scale, for example, diagnosis, or college major. When there are multiple research factors being studied simultaneously, each may relate to Y (and each other) in any of these ways. Thus, when we write $Y = f(C, D, E, \ldots)$, f (as a function of) potentially covers very complex functions, indeed. Yet such complex functions are readily brought under the sway of MRC.

How so? Most readers will know that MRC is often (and properly) referred to as *linear* MRC and may well be under the impression that correlation and regression are restricted to the study of straight-line relationships. This mistaken impression is abetted by the common usage of "linear" to mean "rectilinear," and "nonlinear" to mean "curvilinear." We are thus confounded by what is virtually a pun. What is literally meant by "linear" is any relationship of the form

$$Y = a + bU + cV + dW + eX + \cdots, \tag{1.1.1}$$

where the lower-case letters are constants (either positive or negative) and the capital letters are variables. Y is said to be "linear in the variables U, V, etc." because it is confected by taking certain constant amounts (b, c, etc.) of each variable, and the constant a, and simply adding them together. Were we to proceed in any other way, the resulting function would not be linear in the variables, by definition. But in the fixed-model framework in which we operate, there is no constraint on the nature of the variables. That being the case, they may be chosen so as to define relationships of *any* shape, rectilinear or curvilinear, or of no shape at all (as for unordered nominal scales), and all the complex combinations of these which multiple factors can produce.

Multiple regression equations are, indeed, linear; they are exactly of the form of Eq. (1.1.1). Yet they can be used to describe such complex relationships as the length of psychiatric hospital stay as a function of symptom ratings on admission, diagnosis, age, sex, and average length of prior hospitalizations (if any). This relationship is patently not rectilinear, yet readily described by a linear multiple regression equation.

To be sure, not all or even more relationships studied in the behavioral sciences are of this order of complexity, but the obvious point is that the capacity of MRC to take any degree or type of shape-complexity in its stride is yet another of the important features which make it truly a *general* data-analytic system.

General and Conditional Relationships

Some relationships between Y and some factor C remain the same in regard to degree and form over variation in other factors D, E, F. We will call such relationships *general* or *unconditional*. The definition of a general relationship holds quite apart from how or whether these other factors relate to Y or to C. This might be the case, for example, if Y is a measure of perceptual acuity and C is age. Whatever the form and degree of relationship, if it remains the same under varying conditions of educational level (D), ethnic group (E), and sex (F), then the relationship can be said to be general (insofar as these other factors are concerned). Note that this generality obtains whatever the relationship between acuity and D, E, and F, between age (C) and D, E, F, or among D, E, and F. The *Y-C relationship* can thus be considered unconditional with regard to, or independent of, D, E, and F.

Now consider the same research factors but with Y as a measure of attitude towards racial integration. The form and/or degree of relationship of age to Y is now almost certain to vary as a function of one or more of the other factors: it may be stronger or shaped differently at lower educational levels than higher (D), and/or in one ethnic group than another (E), and/or for men compared to women (F). The relationship of Y to C is now said to be conditional on D and/or E and/or F. In AV contexts, such relationships are called *interactions*, for example, if the *C-Y relationship* is not constant over different values of D, there is said to be a $C \times D$ ("age by educational level") interaction. Greater complexity is possible: the C-Y relationship may be constant over levels of D taken by themselves, and over levels of E taken by themselves, yet may be conditional on the *combination* of D and E levels. Such a circumstance would define a "second-order" interaction, represented as $C \times D \times E$ (with, in this case, neither $C \times D$ nor $C \times E$ present). Interactions of even higher order, and thus even more complex forms of conditionality, are also theoretically possible.

To forestall a frequent source of confusion, we emphasize the fact that the existence of a $C \times D$ interaction is an issue quite separate from the relationship of C to D, or D to Y. However age may relate to education, or education to attitude, the existence of $C \times D$ means that the relationship of age to attitude is conditioned by (depends on, varies with) education. Since such interactions are symmetrical, this would also mean that the relationship of education to attitude is conditioned by age.

One facet of the complexity of the behavioral sciences is the frequency with which conditional relationships are encountered. Relationships among variables often change with changes in experimental conditions (treatments, instructions, experimental assistants, etc.), age, sex, social class, ethnicity, diagnosis, religion, geographic area, etc. Causal interpretation of such conditional relationships is even more difficult than it is for general relationships, but it is patently important that conditionality be detected when it exists.

Conditional relationships may be studied directly, but crudely, by partitioning the data into subgroups on the conditioning variable, determining the relationship in each subgroup, and comparing them. However, problems of small sample size and difficulties in the statistical comparison of measures of relationship from subgroup to subgroup are likely to arise. Factorial design AV provides for assessing conditional relationships (interactions), but is constrained to research factors in nominal form and becomes awkward when the research factors are not orthogonal. The versatility of the MRC system obtains here—conditional relationships of any order of complexity, involving research factors with information in any form, and either correlated or uncorrelated, can be routinely handled without difficulty. (See Chapter 8.)

In summary, the generality of the MRC system of data analysis appropriately complements the complexity of the behavioral sciences, where "complexity" is intended to convey simultaneously the ideas of multiplicity and correlation among potential causal influences, the variety of forms in which information is couched, and in the shape and conditionality of relationships. Multiple regression/correlation also provides a full yield of measures of "effect size" with which to quantify various aspects of relationships (proportions of variance, correlation and regression coefficients). Finally, these measures are subject to statistical hypothesis testing, estimation, and power-analytic procedures.

1.1.3 Causal Analysis and Multiple Regression/Correlation

The rapid progress of the natural sciences and their technologies during the last three centuries has largely been due to the evolution of that remarkable invention, the controlled experiment. Control has been achieved by the care and precision of the manipulation of *treatments*, by isolation of the experiment from extraneous influences, and most recently and particularly in the life sciences, by randomization of units to treatments, thus assuring that "all other things are equal." The virtue of the experiment lies in the simplicity of its causal model: with manipulative control of the treatment, randomization assures that the output is a direct causal consequence of the treatment and not of other causes residing in initial differences between treatment groups.

The experimental paradigm has served the behavioral and social sciences well in those areas wherein its demands can be met, for example, in the traditional area of experimental psychology (learning, memory, perception), physiological psychology, comparative psychology, some aspects of social psychology, and the technologies related to these fields. Practical considerations have limited its utility in some other fields (e.g., education, clinical psychology, program evaluation, psychiatric epidemiology, industrial organization theory), and its application is effectively impossible in sociology, economics, political science, and anthropology. Either the putative causes or effects can not be produced by investigators (e.g., schizophrenia, low income, fascism, high gross national product, egalitarian management structures), or randomization is precluded, or both.

Where experimentation is not possible, scientists have been forced to develop theories from their passive observation of phenomena. In the human sciences, the phenomena are highly variable, putative causes are many and obscure, effects often subtle and their manifestations delayed, and measurement is difficult. Small wonder that progress in these fields has been slow.

Enter the analysis of causal models (path analysis, structural equation systems). Originating in genetics (Wright, 1921) and econometrics about a half century ago, and proceeding independently, there has developed a coherent scheme for the quantitative analysis and testing of theories based on the observation of phenomena as they occur naturally. Beginning in the 1960s, these methods grew in sociology and related fields (Blalock, 1971; Goldberger & Duncan, 1973), and in the 1970s in education and psychology (Kenny, 1979).

The basic strategy of the analysis of causal models is first to state a theory in terms of the variables that are involved and, quite explicitly, of what causes what and what does not, usually aided by causal diagrams. The observational data are then employed to determine whether the causal model is consistent with them, and estimate the strength of the causal parameters. Failure of the model to fit the data results in its falsification, while a good fit allows the model to survive, but not be proven, since other models might provide equal or better fits.

Causal model analysis provides a formal calculus of inference which promises to be as important to the systematically observing scientist as is the paradigm of the controlled experiment to the systematically experimenting scientist. Although still new and in the process of rapid development, it seems clear to us that nonexperimental inference that is not consistent with its fundamental principles is simply invalid.

Now, the major analytic tool of causal models analysis is MRC, and particularly regression analysis. Even the simplest regression equation that states that Y is a linear function of X carries, in its asymmetry, the implication that X causes Y, and not the other way around. As the number of variables causing Y increases, we enter the realm of ***multiple*** regression analysis. As the complexity of the causal model increases, we develop systems of interlocking regression analysis in which a variable may be a cause in one regression equation and an effect in another. Further complexity (i.e., reciprocal causality) may require that we change our methods of estimating causal parameters, but our basic tool remains the regression equation.

We find the old saw that " correlation does not mean causation," although well intentioned, to be grossly misleading. Causation manifests itself in correlation, and its analysis can only proceed through the systematic analysis of correlation and regression. From the very beginning of the presentation of MRC methods in the next chapter, our exposition is informed by the concepts of causal analysis. After the basic devices of MRC are presented, an entire chapter is devoted to causal analysis and its exploitation of these devices in practical working methods. We hope this material serves as an introduction to this most important topic in research methodology in the behavioral and social sciences.

1.2 ORIENTATION

This book was written to serve as a textbook and manual in the application of the MRC system for data analysis by students and practitioners in the diverse areas of inquiry of the behavioral sciences. As its authors, we had to make many decisions about the level, breadth, emphasis, tone, and style of exposition. Its readers may find it useful, at the outset, to have our orientation and the basis for these decisions set forth.

1.2.1 Approach

Nonmathematical

Our presentation of MRC is nonmathematical. Of course, MRC is itself a product of mathematical statistics, based on matrix algebra, the calculus, and probability theory—branches of mathematics familiar only to math majors. There is little question that such a background makes possible a level of insight otherwise difficult to achieve. However, since it is only infrequently found in behavioral scientists, it is bootless to proceed on such a basis, however desirable it may be in theory. Nor do we believe it worthwhile, as is done in some statistical textbooks addressed to this audience, to attempt to provide the necessary mathematical background in condensed form in an introductory chapter or two and then proceed as if it were a functioning part of the reader's intellectual equipment. In our experience, that simply does not work—it serves more as a sop to the author's conscience than as an aid to the reader's comprehension.

We thus abjure mathematical proofs, as well as unnecessary offhand references to mathematical concepts and methods not likely to be understood by the bulk of our audience. In their place, we heavily emphasize detailed and deliberately redundant verbal exposition of concrete examples drawn from the behavioral sciences. Our experience in teaching and consulting convinces us that our audience is richly endowed in the verbal, logical, intuitive kind of intelligence that makes it possible to understand how the MRC system works, and thus use it effectively. (Dorothy Parker said, "Flattery will get you anywhere.") This kind of understanding is eminently satisfactory (as well as satisfying), since it makes possible the effective use of the system. We note that to drive a car, one does not need to be a physicist, nor an automotive engineer, nor even an auto mechanic, although some of the latter's skills are useful when you are stuck on the highway, and that is the level we aim for.

Flat assertions, however, provide little intellectual nourishment. We seek to make up for the absence of mathematical proofs by providing demonstrations instead. For example, the regression coefficient for a dichotomous or binary (male-female, yes-no) independent variable (scored 0–1) equals the difference between the two groups' Y means. Instead of offering the six or seven lines of algebra that would constitute a mathematical proof, we demonstrate that it holds, using a small set of data. True, this proves nothing, since the result may be accidental, but the curious reader can check it out on his own data (and we urge that such checks be made throughout). Whether it is checked out or not, however, we believe that most of our audience would profit more from the demonstration than the proof. If the absence of proof bothers some Missourians, all we can do is pledge our good faith.

Applied

The first word in this book's title is "applied." The heavy stress on illustrations serves not only the function of clarifying and demonstrating the abstract principles being taught, but also that of exemplifying the kinds of applications possible, that is, providing working models. We attend to theory only insofar as sound application makes necessary. The emphasis is on "how to do it." This opens us to the contemptuous charge of writing a "cookbook," a charge we deny, since we do not neglect the whys and wherefores. If the charge is nevertheless pressed, we can only add the observation that in the kitchen, cookbooks are likely to be found more useful than textbooks in organic chemistry.

Data-Analytic

The mathematical statistician proceeds from exactly specified premises (independent random sampling, normality of distribution, homogeneity of variance), and by the exercise of his ingenuity and appropriate mathematical theory, arrives at exact and necessary consequences (F distribution, statistical power functions). He is, of course, fully aware of the fact that no set of real data will exactly conform to the formal premises from which he starts, but this is not properly his responsibility. As all mathematicians, he works with abstractions to produce formal models whose "truth" lies in their self-consistency. Borrowing their language, we might say that inequalities are symmetrical: just as behavioral scientists are not mathematicians, mathematicians are not behavioral scientists.

The behavioral scientist relies very heavily on the fruits of the labors of theoretical statisticians. They provide guides for teasing out meaning from data, limits on inference, discipline in speculation. Unfortunately, in the textbooks addressed to behavioral scientists, statistical methods have often been presented more as harsh straightjackets or Procrustean beds than as benign reference frameworks. Typically, a method is presented with some emphasis on its formal assumptions. Readers are advised that the failure of a set of data to meet these assumptions renders the method invalid. All too often, the discussion ends at this point. Presumably, the offending data are to be thrown away.

Now this is, of course, a perfectly ridiculous idea from the point of view of working scientists. Their task is to contrive situations that yield information about substantive scientific issues—*they must and will analyze their data*. In doing so, they will bring to bear, in addition to the tools of statistical analysis, their knowledge of theory, past experience with similar data, hunches, and good sense, both common and uncommon. They would rather risk analyzing their data incorrectly than not at all. For them, data analysis is not an end in itself, but the next-to-last step in a sequence which culminates in providing information about the phenomena. This is by no means to say that they need not be painstaking in their efforts to generate and perform analyses of data from which unambiguous conclusions may be drawn. But they must translate these efforts into substantive information.

Most happily, the distinction between "data analysis" and "statistical analysis" has been made and given both rationale and respectability by one of the world's foremost mathematical statisticians, John Tukey. In his seminal *The Future of Data Analysis* (1962), Tukey describes data analysis as the special province of scientists with substantial interest in methodology. Data analysts employ statistical analysis as the most important tool in their craft, but they employ it together with other tools, and in a spirit quite different from that which has come to be associated with it from its origins in mathematical statistics. Data analysis accepts "inadequate" data, and is thus prepared to settle for "indications" rather than "conclusions." It risks a greater frequency of errors in the interest of a greater frequency of occasions when the right answer is "*suggested*." It compensates for cutting some statistical corners by using scientific as well as mathematical judgment, and by relying upon self-consistency and repetition of results. Data analysis operates like a detective searching for clues rather than like a bookkeeper seeking to prove out a balance. In describing data analysis, Tukey has provided insight and rationale into the way good scientists have always related to data.

The spirit of this book is strongly data-analytic, in exactly the above sense. We recognize the limits on inference placed by the failure of real data to meet some of the formal assumptions which underly fixed-model MRC, but are disposed to treat the limits as broad rather than narrow. We justify this by mustering whatever technical evidence there is in the statistical literature (for example, on the "robustness" of statistical tests), and by drawing upon our own and others' practical experience, even upon our intuition, all in the interest of getting on with the task of making data yield their meaning. If we risk error, we are more than compensated by having a system of data analysis that is general, sensitive, and fully capable of reflecting the complexity of the behavioral sciences and thus of meeting the needs of behavioral scientists.

1.2.2 Computation, the Computer, and Numerical Results

Computation

Like all mathematical procedures involving simultaneous attention to multiple variables, MRC makes large computational demands. As the size of the problem increases, the amount of computation required increases enormously; for example, the computational time required on a desk calculator for a problem with $k = 10$ independent variables and $n = 400$ cases is measured in days! With such prodigious amounts of hand calculation, the probability of coming through the process without serious blunders (misreading values, inversion of digits, incorrect substitution, etc.) cannot be far from zero. Rigorous checking procedures can assure accuracy, but at the cost of increasing computational man-days. The only solution is *not* to do the calculation on a desk calculator.

An important reason for the rapid increase during the past two decades in the use of multivariate[4] statistical procedures generally, and for the emergence of MRC as a general data-analytic system in particular, is the computer revolution. During this period, computers have become faster to a degree that strains comprehension, more "user oriented," and, most important of all, more widely available. Computer facilities are increasingly looked upon as being as necessary in academic and scientific settings as are library facilities. And progressive simplification in their utilization ("user orientation") makes the necessary know-how fairly easy to acquire. Fundamentally, then, we assume that MRC computation, in general, will be accomplished by computer.

Early in the book, in our exposition of bivariate correlation and regression and MRC with two independent variables, we give the necessary details with worked examples for calculation by desk or pocket calculators (or, in principle, pencil and yellow pad). This is done because the intimate association with the arithmetic details makes plain to the reader the nature of the process: exactly what is being done, with what purpose, and to what result. With two or three variables, where the computation is easy, not only can one see the fundamentals, but there is laid down a basis for generalization to many variables, where the computational demands are great.

With k independent and one dependent variable, MRC computation requires, *to begin with*, $k + 1$ means and standard deviations, and the matrix of $k(k + 1)/2$ correlation coefficients between all pairs of $k + 1$ variables. It is at this point that the serious computation begins, that is, the solution of k simultaneous equations, most readily accomplished by a laborious matrix-arithmetic procedure called

[4]Usage of the term *multivariate* varies. Some authors restrict it to procedures where multiple *dependent* variables are used, by which definition MRC would not be included. However, increasingly and particularly among applied statisticians and behavioral scientists, the term is used to cover all statistical applications wherein "multiple variates are considered in combination" (Cooley & Lohnes, 1971, p. 3), either as dependent or independent variables, or both, or, as in factor analysis, neither; see Tatsuoko (1971) and Van de Geer (1971). See Chapter 12 and Appendix 4 for a consideration of MRC in relationship to other multivariate methods.

"inversion." Appendix 1 describes the mathematical basis of MRC including the role of the centrally important operation of matrix inversion, and the content and meaning of the elements of the inverse matrix. In Appendix 2, we give the actual computational (arithmetic) operations of matrix inversion and multiplication for MRC, suitable for use with a desk or pocket calculator. Although, in principle, the computational scheme given in Appendix 2 may be used with any number of variables, it becomes quite time consuming, and rapidly more onerous, as k increases beyond five or so.[5] The reader without access to computers has our sympathy, but he can manage the computation in small problems by following the procedure in Appendix 2, and may even be rewarded by insights which may accrue from this more intimate contact with the analysis.

But "the human use of human beings" does not include days spent at a desk calculator. As we have noted, we primarily rely on computers for MRC computation. Most of our readers will have access to a computer, and will either have, or be able quickly to obtain, the modest know-how needed to use the available "canned" programs for MRC. Appendix 3 is devoted to this topic, and includes a description of the characteristics of the most popular and widely available programs, and of the considerations which should enter into one's choice among them. It should be consulted in conjunction with a trip to the computer laboratory to investigate what is available.

We expect that our readers will find the material on "heavy" computation useful, but we have deliberately placed it outside the body of the text to keep it from distracting attention from our central emphasis, which is on understanding how the MRC system works, so that it may be effectively used in the exploitation of research data.

Our attitude toward computing as such explains the absence of chapter-end problems for the reader to work. Some of the purposes of such exercises can be achieved by carefully following through the details of the many worked illustrative examples in the body of the text. But the highest order of understanding is to be attained when readers apply the methods of each chapter *to data of their own*, or data with which they are otherwise familiar. There is no more powerful synergism for insight than the application of unfamiliar methods to familiar data.

Numerical Results: Reporting and Rounding

With minor exceptions, the computation of the illustrative problems which fill this book was all accomplished by computer, using various programs and different computers. The numerical results carried in the computer are accurate to at least six significant figures and are printed out to at least four (or four decimal places). We see little point to presenting numerical results to as many places as

[5]It is difficult to set a value here—who is to say what another will find computationally onerous? Some people find balancing their checkbook a nightmare; others actually enjoy large quantities of arithmetic, particularly when it involves their own data. Five seems a reasonable compromise. But it should be kept in mind that computation of time increases roughly (no pun intended) with k^5 (Darlington & Boyce, 1982)!

the computer may provide, since the resulting "accuracy" holds only for the sample data analyzed, and, given the usual level of sampling error, is quite meaningless vis-à-vis the values of the population parameters. We mean nothing more complicated than the proposition, for example, that when the computer mindlessly tells us that, in a sample of the usual size, the product moment correlation between X and Y is .34617952, a guaranteed accurate result, at least the last five digits could be replaced by random numbers with no loss.

In this book, we generally follow the practice of reporting computed correlation and regression coefficients of all kinds and significance test results rounded to three places (or significant figures), and squared correlations (proportions of variance) rounded to four. (Occasional departures from this practice are for specific reasons of expository clarity or emphasis.) Thus, the above r would be reported as .346. But the computer treats it in the calculations in which it is involved as .34617952 . . . , and its square as $.34617952\ldots^2$. Thus, when we have occasion to report the square of this r, we do not report $.346^2$, which equals .1197 when rounded to four places, but $.34617952\ldots^2$ rounded to four places, which is .1198. When the reader tries to follow our computations (which he should), he will run across such apparent errors as $.346^2 = .1198$ and others which are consequent on his use in computation of the reported three-digit rounded values. These are, of course, not errors at all, but inevitable rounding discrepancies. Checks which agree within a few points in the fourth decimal place may thus be taken as correct.

Significance Test Results and the Appendix Tables

We employ classical null hypothesis testing, in which the probability of the sample result, P, is compared to a prespecified significance criterion α. If $P <$ (is less than) α, the null hypothesis (usually that analogous population value is zero) is rejected, and the sample result is deemed statistically "significant" at the α level. In the tests we predominantly use (F and t), the actual value of P is not determined.[6] Instead, the F or t value for the sample result is computed by the appropriate formula, and the result is compared with the value of F or t at the α criterion value found from a table in the Appendix. Then, if the sample value exceeds the criterion value, we conclude that $P < \alpha$, *and the null hypothesis is rejected.*

We make provision in the Appendix Tables of F and t, and in those used for statistical power analysis, for the significance criteria $\alpha = .01$ and $\alpha = .05$. We see no serious need in routine work for other α values. The $\alpha = .05$ criterion is so widely used as a standard in the behavioral sciences that it has come to be understood to govern when a result is said to be statistically significant in the absence of a specified α value. The more stringent $\alpha = .01$ criterion is used by some investigators routinely as a matter of taste or of tradition in their research

[6]That is, not by us in this book. Most computer programs compute and print out the actual P for each F or t given (see Appendix 3).

area, by others selectively when they believe the higher standard is required for substantive or structural reasons. We are inclined to recommend its use in research involving many variables and hence many hypothesis tests as a control on the incidence of spuriously significant results. The choice of α also depends importantly on considerations of statistical power (the probability of rejecting the null hypothesis), which is discussed in several places and in most detail in Section 4.5.

In reporting the results of significance tests for the many worked examples, we follow the general practice of attaching double asterisks to an F or t value to signify that $P < .01$, and a single asterisk to signify that $P < .05$ (but not .01). No asterisk means that the F or t is not significant, that is, P exceeds .05.

The statistical tables in the Appendix were largely abridged from Owen (1962) and from Cohen (1977), with some values computed by us. The entry values were selected so as to be optimally useful over a wide range of MRC applications. For example, we provide for many values of numerator degrees of freedom (numbers of independent variables) in the F and L tables, and similarly for denominator (error) degrees of freedom in the F and t tables and for n in the power tables for r. On the other hand, we do not cover very low values for n, since they are almost never used. The coverage is sufficiently dense to preclude the need for interpolation in most problems; where needed, linear interpolation is sufficiently accurate for almost all purposes. On very rare occasions more extensive tables may be required, for which Owen (1962) is recommended.

1.2.3 The Spectrum of Behavioral Science

When we address behavioral scientists, we are faced with an exceedingly heterogeneous audience, indeed. We note in passing that our intended audience ranges from student to experienced investigator, and from possession of modest to fairly advanced knowledge of statistical methods. With this in mind, we assume a minimum background for the basic exposition of the MRC system, but at some later points and infrequently, we must make some assumptions about background which may not hold for some of our readers, in order that we may usefully address some others. Even then, we try hard to keep everyone on board.

But it is with regard to substantive interests and investigative methods and materials that behavioral scientists are of truly mind-boggling diversity. The rubric "behavioral science" has no exactly delimited reference, but we use it broadly, so that it covers the "social," "human," and even "life" sciences, everything from the physiology of behavior to cultural anthropology, in both their "pure" and "applied" aspects. Were it not for the fact that the methodology of science is inherently more general than its substance, a book of this kind would not be possible. However, our target audience is made up, not of methodologists, but of people whose primary interests lie in a bewildering variety of fields.

We have sought to accommodate to this diversity, even to capitalize upon it. Our illustrative examples are drawn from different areas, assuring the comfort of familiarity for most of our readers at least some of the time. They have been composed with certain ideas in mind: their content is at a level which makes them intellectually accessible to nonspecialists, and they are all fictitious, so they can accomplish their illustrative purposes efficiently and without the distractions and demands for specialized knowledge which would characterize real data. We try to use the discussion of the examples in a way which may promote some cross-fertilization between fields of inquiry—when discussed nontechnically, some problems in a given field turn out to be freshly illuminated by concepts and approaches from other fields. We may even contribute to breaking down some of the methodological stereotypy to be found in some areas, where data are analyzed traditionally, rather than optimally.

1.3 PLAN

1.3.1 Content

The first part of this book (Chapters 1 through 4) develops the basic ideas and methods of multiple correlation and regression. Chapter 2 treats simple linear correlation for two variables, X and Y, and the related linear regression model, with Y as a dependent variable and X as a single independent variable, in both their descriptive and inferential (statistical hypothesis testing and power analysis) aspects. In the first part of Chapter 3, the MRC model is extended to two independent variables, which introduces the important ideas of multiple and partial regression and correlation, and the distinction between simultaneous and hierarchical MRC. The relevance to causal models is shown. In the latter part of Chapter 3, the conceptually straightforward generalization from two to k independent variables is made.

Up to this point, for the most part, the treatment is fairly conventional. Chapter 4, however, introduces a further generalization of MRC, wherein the independent variables are organized into h *sets*, each made up of one or more variables. The utility of this extension arises from the fact that the research factors (presumed causes) being studied are expressed, in general, as sets, to which the ideas of partial relationship and of hierarchical versus simultaneous analyses are applied. Simple methods for hypothesis testing and statistical power analysis are included. With this chapter, the basic structure of MRC as a general data-analytic system is complete.

The stage having been set, Part II proceeds to detail in a series of chapters how information in any form may be represented as sets of independent variables, and how the resultant MRC yield for sets and their constituent variables is interpreted. In Chapter 5, various methods are described for representing nominal scales (for example, experimental treatment, diagnosis, religion), and the opportunity is grasped to show that the MRC results include those produced by AV, and more. Chapter 6 performs the same task for quantitative (ratio, interval, ordinal) scales, showing how various methods of representation may be used to determine the presence and form of curvilinear relationship with Y. Chapter 7 is concerned with the representation of missing data, a ubiquitous problem in the behavioral sciences, and shows how this property of a research factor may be used as positive information. Chapter 8 presents and generalizes the idea of interactions as conditional relationships, and shows how interactions among research factors of any type may be simply represented as sets, incorporated in analyses, and interpreted.

In Part III ("Applications"), Chapter 9 presents an introductory treatment of causal models. Utilizing causal diagrams and the basic MRC methods detailed in the earlier chapters, it provides working methods for the analysis of recursive causal models, and points the direction in which more complex models involving reciprocal causality may be analyzed.

Chapter 10 provides a culmination of the ideas about multiple sets (Chapter 4), implemented by the methods of representing research factors (Chapters 5 through 7) and their interactions (Chapter 8). It shows how conventional ACV may be accomplished by MRC, and then greatly generalizes ACV, first to accommodate multiple, nonlinear, and nominal covariate sets, and then to extend to quantitative research factors. The nature of partialling is closely scrutinized, and the problem of fallible (unreliable) covariates is addressed, as is the use of the system in the study of change.

Chapter 11 extends MRC analysis to repeated measurement and matched-subject research designs. In Chapter 12, canonical correlation analysis and other multivariate methods are surveyed and related to MRC analysis, from which some novel analytic methods emerge.

In the Appendices, we provide the mathematical background for MRC (Appendix 1), the hand computation for k variables together with the interpretation of the results of matrix inversion (Appendix 2), and a discussion of the use of available computer programs to accomplish MRC analyses (Appendix 3). Appendix 4 presents a new data-analytic method called *set correlation*. This is a multivariate method which generalizes MRC to include sets (or partialled sets) of dependent variables and in so doing, generalizes multivariate methods and yields novel data-analytic forms. Finally, the necessary statistical tables are provided.

For a more detailed synopsis of the book's contents, the reader is referred to the summaries at the ends of the chapters.

1.3.2 Structure: Numbering of Sections, Tables, and Equations

Each chapter is divided into major sections, identified by the chapter and section numbers, for example, Section 5.3 ("Dummy-Variable Coding") is the third major section of Chapter 5. The next lower order of division within each major section is further suffixed by its number, for example, Section 5.3.4 ("Dummy-Variable Multiple Regression/Correlation and Analysis of Variance") is the fourth subsection of Section 5.3. Further subdivisions are not numbered, but titled with an italicized heading.

Tables, figures, and equations within the body of the text are numbered consecutively within major sections. Thus, for example, Table 5.3.4 is the fourth table in Section 5.3, and Eq. (5.3.4) is the fourth equation in Section 5.3. (We follow the usual convention of giving equation numbers in parenthesis.) A similar plan is followed in the four Appendices. The reference statistical tables make up a separate appendix and are designated by letters as Appendix Tables A through F.

1.4 SUMMARY

This introductory chapter begins with an overview of MRC as a data-analytic system, emphasizing its generality and superordinate relationship to the analysis of variance/covariance. Multiple regression/correlation is shown to be peculiarly appropriate for the behavioral sciences in its capacity to accommodate the various types of complexity which characterize them: the multiplicity and correlation among causal influences, the varieties of form of information and shape of relationship, and the frequent incidence of conditional (interactive) relationships. The special relevance of MRC to the formal analysis of causal models is described. (Section 1.1)

The book's exposition of MRC is nonmathematical, and stresses informed application to scientific and technological problems in the behavioral sciences. Its orientation is "data-analytic" rather than statistical-analytic, an important distinction that is discussed. Concrete illustrative examples are heavily relied upon. The means of coping with the computational demands of MRC by desk calculator and computer are briefly described and the details largely relegated to appendices so as not to distract the reader's attention from the conceptual issues. The ground rules for reporting numerical results (including a warning about rounding discrepancies) and those of significance tests are given, and the statistical tables in the appendix are described. Finally, we acknowledge the heterogeneity of background and substantive interests of our intended audience, and discuss how we try to accommodate to it and even exploit it to pedagogical advantage. (Section 1.2)

The chapter ends with a brief outline of the book, and the scheme by which sections, tables, and equations are numbered. (Section 1.3)

References

Acton, F. S. *Analysis of straight-line data*. New York: Wiley, 1959.

Althauser, R. P. Multicollinearity and non-additive regression models. In H. M. Blalock, Jr. (Ed.), *Causal models in the social sciences*. Chicago: Aldine-Atherton, 1971.

Anderson, R. L., & Houseman, E. E. *Tables of orthogonal polynomial values extended to N = 104* (Research bulletin No. 297). Ames, Iowa: Iowa Agricultural Experiment Station, 1942.

Anscombe, F. J., & Tukey, J. W. The examination and analysis of residuals. *Technometrics*, 1963, *5*, 141–160.

Baker, B., Hardyck, C. D., & Petrinovich, L. F. Weak measurements vs. strong statistics: An empirical critique of S. S. Stevens' proscriptions on statistics. *Educational and Psychological Measurement*, 1966, *26*, 291–309.

Bartlett, M. S. Square-root transformation in analysis of variance. *Journal of the Royal Statistical Society Supplement*, 1936, *3*, 68–78.

Bentler, P. Multivariate analysis with latent variables: Causal modeling. *Annual Review of Psychology*, 1980, *31*, 419–456.

Binder, A. M. Considerations of the place of assumptions in correlational analysis. *American Psychologist*, 1959, *14*, 504–510.

Blalock, H. M., Jr. *Causal inferences in nonexperimental research*. Chapel Hill: University of North Carolina Press, 1964.

Blalock, H. M., Jr. (Ed.). *Causal models in the social sciences*. Chicago: Aldine-Atherton, 1971.

Blalock, H. M., Jr., & Blalock, A. B. (Eds.). *Methodology in social research*. New York: McGraw–Hill, 1968.

Bock, R. D. *Multivariate statistical methods in behavioral research*. New York: McGraw–Hill, 1975.

Boneau, C. A. The effects of violations of assumptions underlying the *t* test. *Psychological Bulletin*, 1960, *57*, 49–64.

Bottenberg, R. A., & Ward, J. H., Jr. *Applied multiple linear regression*. (PRL-TRD-63-6). Lackland Air Force Base, Texas, 1963. (NTIS No. AD-413128).

Budescu, D. V. A note on polynomial regression. *Multivariate Behavioral Research*, 1980, *15*, 497–508.

Campbell, D. T., & Erlebacher, A. How regression artifacts in quasi-experimental evaluations can mistakenly make compensatory education look harmful. In J. Hellmuth (Ed.), *Disadvantaged child, Vol. 3: Compensatory education: A national debate*. New York: Brunner/Mazel, 1970.

Campbell, D. T., & Stanley, J. C. *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally, 1966.

Carmer, S. G., & Swanson, M. R. An evaluation of ten pairwise multiple comparison procedures by Monte Carlo methods. *Journal of the American Statistical Association*, 1973, *68*, 66–74.

Cattin, P. The predictive power of ridge regression: some quasi-simulation results. *Journal of Applied Psychology*, 1981, *66*, 282–290.

Chesire, L., Safir, M., & Thurstone, L. L. *Computing diagrams for the tetrachoric correlation coefficient*. Chicago: University of Chicago Bookstore, 1933.

Cliff, N. What is and isn't measurement. In G. Keren (Ed.), *Statistical and methodological issues in psychology and social science research*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1982.

Cliff, N., & Krus, D. J. Interpretation of canonical analysis: Rotated versus unrotated solutions. *Psychometrika*, 1976, *41*, 35–42.

Cochran, W. G. Some consequences when the assumptions for the analysis of variance are not satisfied. *Biometrics*, 1947, *3*, 22–38.

Cohen, J. The statistical power of abnormal-social psychological research: A review. *Journal of Abnormal and Social Psychology*, 1962, *65*, 145–153.

Cohen, J. Some statistical issues in psychological research. In B. B. Wolman (Ed.), *Handbook of clinical psychology*, New York: McGraw–Hill, 1965.

Cohen, J. Multiple regression as a general data-analytic system. *Psychological Bulletin*, 1968, *70*, 426–443.

Cohen, J. Causation → Correlation. *Contemporary Psychology*, 1973, *18*, 453–455. (a)

Cohen, J. Statistical power analysis and research results. *American Educational Research Journal*, 1973, *10*, 225–229. (b)

Cohen, J. *Statistical power analysis for the behavioral sciences* (Rev. ed.). New York: Academic Press, 1977.

Cohen, J. Partialed products *are* interactions; Partialed powers *are* curve components. *Psychological Bulletin*, 1978, *85*, 858–866.

Cohen, J. Trend analysis the easy way. *Educational and Psychological Measurement*, 1980, *40*, 565–568.

Cohen, J. Set correlation as a general multivariate data-analytic method. *Multivariate Behavioral Research*, 1982, *17*, 301–341.

Cohen, J., & Nee, J. CORSET: A FORTRAN program for set correlation. *Educational and Psychological Measurement*, 1983, 1, No. 1.

Cohen, P., Belmont, L., Dryfoos, J., Stein, Z., & Zayac, S. The effects of teenaged motherhood and maternal age on offspring intelligence. *Social Biology*, 1980, *27*, 138–154.

Coleman, J. S. *Introduction to mathematical sociology*. Glencoe, Ill.: Free Press, 1964.

Cook, T. D., & Campbell, D. T. *Quasi-experimentation. Design & analysis issues for field settings*. Chicago: Rand McNally, 1979.

Cooley, W. W., & Lohnes, P. R. *Multivariate data analysis*. New York: Wiley, 1971.

Cramer, E. M., & Nicewander, W. A. Some symmetric, invariant measures of multivariate association. *Psychometrika*, 1979, *44*, 43–54.

Cronbach, L. J. *Essentials of psychological testing* (3rd ed.). New York: Harper & Row, 1970.

Cronbach, L. J., & Furby, L. How we should measure "change"—or should we? *Psychological Bulletin*, 1970, *74*, 68–80.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley, 1972.

Darlington, R. B. Reduced variance regression. *Psychological Bulletin*, 1978, *85*, 1238–1255.

Darlington, R. B., & Boyce, C. M. Ridge and other new varieties of regression. In G. Keren (Ed.), *Statistical and methodological issues in psychology and social sciences research*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1982.

Dawes, R. M. The robust beauty of improper linear models in decision making. *American Psychologist*, 1979, *34*, 571–582.

Dawes, R. M., & Corrigan, B. Linear models in decision making. *Psychological Bulletin*, 1974, *81*, 95–106.

DeGracie, J. S., & Fuller, W. A. Estimation of the slope and analysis of covariance when the concomitant variable is measured with error. *Journal of the American Statistical Association*, 1972, *67*, 930–937.

Dempster, A. P., Schatzoff, M., & Wermuth, N. A simulation study of alternatives to ordinary least squares. *Journal of the American Statistical Association*, 1977, *72*, 77–91.

Dixon, W. J. (Ed.). *BMDP statistical software: 1981 edition*. Berkely: University of California Press, 1981.

Doll, R., & Peto, R. *Causes of cancer*. Oxford: Oxford Medical Publications, 1981.

Draper, N. R., & Smith, H. *Applied regression analysis* (2nd ed.). New York: Wiley, 1981.

Duncan, O. T. *Introduction to structural equation models*. New York: Academic Press, 1975.

Dunivant, N. *The effects of measurement error on statistical models for analyzing change*. Final report, Grant NIE–G–78–0071, National Institute of Education, U.S. Department of Education, Washington, D.C., November 1981.

Edwards, A. E. *Experimental design in psychological research* (4th ed.). New York: Holt, Rinehart & Winston, 1972.

Ezekiel, M., & Fox, K. A. *Methods of correlation and regression analysis* (3rd ed.). New York: Wiley, 1959.

Fisher, R. A., & Yates, F. *Statistical tables for biological, agricultural and medical research* (6th ed.). New York: Hafner, 1963.

Fornell, C. External single-set components analysis of multiple criterion/multiple predictor variables. *Multivariate Behavioral Research*, 1979, *14*, 323–338.

Freeman, M. F., & Tukey, J. W. Transformations related to the angular and the square root. *Annals of Mathematical Statistics*, 1950, *21*, 607–611.

Fuller, W. A. Regression analysis for sample survey. *Sankhya*, 1975, *37C*, 117–132.

Fuller, W. A., & Hidiroglou, M. A. Regression estimation after correlation for attenuation. *Journal of the American Statistical Association*, 1978, *73*, 99–104.

Games, P. A. Multiple comparisons of means. *American Educational Research Journal*, 1971, *8*, 531–565.

Glass, G. V, Willson, V. K., & Gottman, J. M. *The design and analysis of time-series experiments*. Boulder, Col.: Colorado Associated University Press, 1975.

Goldberger, A. S. *Econometric theory*. New York: Wiley, 1964.

Goldberger, A. S., & Duncan, O. D. (Eds.). *Structural equation models in the social sciences*. New York: Seminar Press, 1973.

Green, B. F., Jr. Parameter sensitivity in multivariate methods. *Multivariate Behavioral Research*, 1977, *12*, 263–288.

Hanushek, E. A., & Jackson, J. E. *Statistical methods for social scientists*. New York: Academic Press, 1977.

Harris, C. W. (Ed.). *Problems in measuring change*. Madison: University of Wisconsin Press, 1963.

Harris, R. J. *A primer of multivariate statistics*. New York: Academic Press, 1975.

Havlicek, L. L., & Peterson, N. L. Effect of the violation of assumptions upon significance levels of the Pearson r. *Psychological Bulletin*, 1977, *84*, 373–377.

Hays, W. L. *Statistics* (3rd ed.). New York: Holt, Rinehart & Winston, 1981.

Heise, D. R. *Causal analysis*. New York: Wiley–Interscience, 1975.

Hooper, J. W. Simultaneous equations and canonical correlation theory. *Econometrica*, 1959, *27*, 245–256.

Hooper, J. W. Partial trace correlations. *Econometrica*, 1962, *30*, 324–331.

Hotelling, H. Relations between two sets of variates. *Biometrika*, 1936, *28*, 321–377.

Hull, C. H., & Nie, N. H. *SPSS update*. New York: McGraw–Hill, 1979.

James, L. R. A test for asymmetric relationships between two reciprocally related variables. *Multivariate Behavioral Research*, 1981, *16*, 63–82.

James, L. R., & Singh, B. K. An introduction to the logic, assumptions, and basic analytic procedures of two-stage least squares. *Psychological Bulletin*, 1978, *85*, 1104–1122.

Johnston, J. *Econometric methods*. New York: McGraw–Hill, 1963.

Jones, M. B. Moderated regression and equal opportunity. *Educational and Psychological Measurement*, 1973, *33*, 591–602.

Jöreskog, K. G. A general method for analysis of covariance structures. *Biometrika*, 1970, *57*, 239–251.

Jöreskog, K. G., & Sörbom, D. *LISREL V: User's guide*. Chicago: National Educational Resources, 1981.

Kenny, D. A. *Correlation and causality*. New York: Wiley–Interscience, 1979.

Kohn, M. L., & Schooler, C. The effect of the substantive complexity of the job on intellectual flexibility: A longitudinal assessment. *American Journal of Sociology*, 1978, *84*, 24–52.

Lawley, D. N., & Maxwell, A. E. Regression and factor analysis. *Biometrika*, 1973, *60*, 331–338.

Lewis, D. *Quantitative methods in psychology*. New York: McGraw–Hill, 1960.

Li, J. C R. *Statistical inference* (Vol. 2). Ann Arbor, Mich.: Edwards Bros., 1964.

Linn, P. L., & Werts, C. E. Measurement error in regression. In G. Keren (Ed.), *Statistical and methodological issues in psychology and social sciences research*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1982.

Lord, F. M. Statistical adjustment when comparing preexisting groups. *Psychological Bulletin*, 1969, *72*, 336–337.

Lord, F. M. Significance test for a partial correlation corrected for attenuation. *Educational and Psychological Measurement*, 1974, 34, 211–220.

Lord, F. M., & Novick, M. R. *Statistical theories of mental test scores*. Reading, Mass.: Addison–Wesley, 1968.

Luce, R. D., Bush, R. R., & Galanter, E. (Eds.). *Handbook of mathematical psychology*. New York: Wiley, 1963.

McKeon, J. J. Canonical analysis. *Psychometric Monographs*, No. 13, 1964.

McNeil, K. A., Kelly, F. J., & McNeil, J. T. *Testing research hypotheses using multiple linear regression*. Carbondale, Ill.: Southern Illinois University Press, 1975.

Miller, R. G., Jr. *Simultaneous statistical inference*. New York: McGraw Hill, 1966.

Mosteller, F., & Tukey, J. W. *Data analysis and regression*. Reading, Mass.: Addison–Wesley, 1977.

Namboodiri, N. K., Carter, L. F., & Blalock, H. M., Jr. *Applied multivariate analysis and experimental designs*. New York: McGraw–Hill, 1975.

Neter, J., & Wasserman, W. *Applied linear statistical models*. Homewood, Ill.: Richard D. Irwin, 1974.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. *SPSS: Statistical package for the social sciences* (2nd ed.). New York: McGraw–Hill, 1975.

Nunnally, J. C. *Psychometric theory* (2nd ed.). New York: McGraw–Hill, 1978.

Olson, C. L. On choosing a test statistic in multivariate analysis of variance. *Psychological Bulletin*, 1976, *83*, 579–586.

Owen, D. B. *Handbook of statistical tables*. Reading, Mass.: Addison–Wesley, 1962.

Overall, J. E. Calculation of adjusted response frequencies using least square regression methods. *Applied Psychological Measurement*, 1980, *4*, 65–78.

Pearson, E. S., & Hartley, H. O. (Eds.). *Biometrika tables for statisticians* (Vol. 1). Cambridge: Cambridge University Press, 1954.

Pedhazur, E. J. *Multiple regression in behavioral research* (2nd ed.). New York: Holt, Rinehart & Winston, 1982.

Pruzek, R. M., & Fredericks, B. C. Weighting predictors in linear models: Alternatives to least squares and limitations of equal weights. *Psychological Bulletin*, 1978, *85*, 254–266.

Rao, C. R. *Linear statistical inference and its applications* (2nd ed.). New York: Wiley, 1975.

Ray, A. A. (Ed.). *SAS user's guide: Statistics, 1982 edition*. Cary, N.C.: SAS Institute, 1982.

Rogosa, D. Comparing nonparallel regression lines. *Psychological Bulletin*, 1980, *88*, 307–321.

Roy, S. N. *Some aspects of multivariate analysis*. New York: Wiley, 1957.

Rozeboom, W. W. Linear correlations between sets of variables. Psychometrika, 1965, *30*, 57–71.

Rozeboom, W. W. Estimation of cross-validated multiple correlation: A clarification. *Psychological Bulletin*, 1978, *85*, 1348–1351.

Rozeboom, W. W. Sensitivity of linear composites of predictor items to differential item weighting. *Psychometrika*, 1979, *44*, 289–296.

Ryan, T. A. Multiple comparisons in psychological research. *Psychological Bulletin*, 1959, *56*, 26–47.

Srikantan, K. S. Canonical association between nominal measurements. *Journal of the American Statistical Association*, 1970, *65*, 284–292.

Steiger, J. H. Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 1980, *87*, 245–251.

Stephenson, W. *The study of behavior*. Chicago: University of Chicago Press, 1953.

Stevens, S. S. Mathematics, measurement, and psychophysics. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951.

Stevens, S. S. The psychophysics of sensory function. In W. A. Rosenblith (Ed.), *Sensory communication*. New York: Wiley, 1961.

Stewart, D., & Love, W. A general canonical correlation index. *Psychological Bulletin*, 1968, *70*, 160–163.

Stumpf, S. A. A note on handling missing data. *Journal of Management*, 1978, *4*, 65–73.

Survey Research Center. *OSIRIS IV user's manual* (6th ed.). Ann Arbor, Mich.: Institute for Social Research, University of Michigan, 1980.

Tatsuoko, M. M. *Multivariate analysis: Techniques for educational and psychological research*. New York: Wiley, 1971.

Theil, H. *Principles of econometrics*. New York: Wiley, 1971.

Timm, N. H. The estimation of variance–covariance and correlation matrices from incomplete data. *Psychometrika*, 1970, *35*, 417–438.

Timm, N. H. *Multivariate analysis with applications in education and psychology*. Monterey, Calif.: Brooks/Cole, 1975.

Timm, N. H., & Carlson, J. E. Part and bipartial canonical correlation analysis. *Psychometrika*, 1976, *41*, 159–176.

Tukey, J. W. Causation, regression and path analysis. In O. Kempthorne (Ed.), *Statistics and mathematics in biology*. Ames, Iowa: Iowa State College Press, 1954.

Tukey, J. W. The future of data analysis. *Annals of Mathematical Statistics*, 1962, *33*, 1–67.

Van de Geer, J. P. *Introduction to multivariate analysis for the social sciences*. San Francisco: W. H. Freeman, 1971.

Wainer, H. Estimating coefficients in linear models: It don't make no never mind. *Psychological Bulletin*, 1976, *83*, 213–217.

Wainer, H. On the sensitivity of regression and regressors. *Psychological Bulletin*, 1978, *85*, 267–273.

Wampler, R. H. A report on the accuracy of some widely used least squares computer programs. *Journal of the American Statistical Association*, 1970, *65*, 549–565.

Ward, J. H., Jr., & Jennings, E. *Introduction to linear models*. Englewood Cliffs, N.J.: Prentice-Hall, 1973.

Warren, R. D., White, J. K., & Fuller, W. A. An error-in-variables analysis of managerial role performance. *Journal of the American Statistical Association*, 1974, *69*, 886–893.

Wilkinson, L. Tests of significance in stepwise regression. *Psychological Bulletin*, 1979, *86*, 168–174.

Wilks, S. S. Certain generalizations in the analysis of variance. *Biometrika*, 1932, *24*, 471–494.

Wilson, W. A note on the inconsistency inherent in the necessity to perform multiple comparisons. *Psychological Bulletin*, 1962, *59*, 96–300.

Winer, B. J. *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill, 1971.

Wright, S. Correlation and causation. *Journal of Agricultural Research*, 1921, *20*, 557–585.

Page Intentionally Left Blank