

# Contents

## 1 The Forest before the Trees

- 1.0 Why Statistics? 1
  - 1.01 Statistics as a Form of Social Control 1
  - 1.02 Objections to Null Hypothesis Significance Testing 2
  - 1.03 Should Significance Tests be Banned? 3
  - 1.04 Math Modeling's the Ultimate Answer 5
  - 1.05 Some Recent Developments in Univariate Statistics 6
    - 1.0.5.1 The MIDS and FEDs criteria as alternatives to power calculation 7
      - Table 1.1 Fraction of Population Effect Size That Must Be Statistically Significant in Order to Achieve a Given Level of Power for Your Significance Test 9
      - 1.0.5.2 Prior-Information Confidence Intervals (PICIs) 9
- 1.1 Why Multivariate Statistics? 10
  - 1.1.1 Bonferroni Adjustment: An Alternative to Multivariate Statistics. 13
  - 1.1.2 Why Isn't Bonferroni Adjustment Enough? 14
- 1.2 A Heuristic Survey of Statistical Techniques 14
  - Table 1.2 Statistical Techniques 16
  - 1.2.1 Student's  $t$  test 17
  - 1.2.2 One-Way Analysis of Variance 18
  - 1.2.3 Hotelling's  $T^2$  21
    - Example 1.1 Anglo versus Chicano Early Memories* 23
  - 1.2.4 One-Way Multivariate Analysis of Variance 24
    - Example 1.2 Inferring Social Motives from Behavior* 25
  - 1.2.5 Higher Order Analysis of Variance 26
  - 1.2.6 Higher Order Manova 27
    - Example 1.3 Fat, Four-eyed, and Female* 28
  - 1.2.7 Pearson  $r$  and Bivariate Regression 28
  - 1.2.8 Multiple Correlation and Regression 31
    - Example 1.4 Chicano Role Models, GPA, and MRA* 33
  - 1.2.9 Path Analysis 34
  - 1.2.10 Canonical Correlation 35
    - Figure 1.1 Multivariate Analyses of Between-Set Relationships* 36
    - Example 1.5 Television Viewing and Fear of Victimization* 37
  - 1.2.11 Analysis of Covariance 38
  - 1.2.12 Principal Component Analysis 40
  - 1.2.13 Factor Analysis 42
    - Example 1.6 Measuring Perceived Deindividuation* 44
  - 1.2.14 Structural Equation Modeling 44

1.3 Learning to Use Multivariate Statistics	45
1.3.1 A Taxonomy of Linear Combinatons	45
1.3.1.1 Averages of subsets of the measures	45
1.3.1.2 Profiles	47
1.3.1.3 Contrasts	47
1.3.2 Why the Rest of the Book?	51
Quiz 1 See How Much You Know after Reading Just One Chapter!	55
Sample Answers to Quiz 1	56

## 2 Multiple Regression: Predicting One Variable from Many

Data Set 1	58
2.1 The Model	59
2.2 Choosing Weights	62
2.2.1 Least Squares Criterion	62
<b>Table 2.1</b> Multiple Regression Analyses of Data Set 1	66
<b>Table 2.2</b> Data Set 1b: A Presumptuous Data Set	68
2.2.2 Maximum Correlation Criterion	69
2.2.3 The Utility of Matrix Algebra	70
2.2.4 Independence of Irrelevant Parameters	72
2.3 Relating the Sample Equation to the Population Equation	74
<b>Table 2.3</b> Summary of Significance Tests for Multiple Regression	77
2.3.1 $R_x$ versus $S_x$ versus $x'x$ as the Basis for MRA	81
<b>Table 2.4</b> Alternative MRA Formulae	84
2.3.2 Specific Comparisons	84
2.3.3 Illustrating Significance Tests	86
<i>Example 2.1</i> Locus of Control, the CPQ, and Hyperactivity	86
<i>Computer break 2-1: CPQ vs. LOC, CPT-C, CPT-E</i>	89
2.3.4 Stepwise Multiple Regression Analysis	95
<i>Example 2.1</i> Revisited	96
2.4 Computer Programs for Multiple Regression	96
2.4.1 Computer Logic and Organization	97
2.4.2 Sage Advice on Use of Computer Programs	98
2.4.3 Computerized Multiple Regression Analysis	100
2.4.3.1 MATLAB	100
2.4.3.2 SPSS REGRESSION, Syntax Window	101
2.4.3.3 SPSS REGRESSION, Point-and-Click	102
2.4.3.4 SAS PROC REG and PROC RSQUARE	103
2.5 Some General Properties of Covariance Matrices	105
2.6 Measuring the Importance of the Contribution of a Single Variable	107
<b>Table 2.5</b> Measures of Importance in MRA	110

2.7 Anova via MRA	103
<b>Table 2.6</b> Relationship Between MRA and Anova Effects Model	105
<i>Example 2.2 In-Group/Out-Group Stereotypes</i>	106
<b>Table 2.7</b> Coding of MRA Level-Membership Variables for Study of Stereotypes	113
<i>Example 2.3 Negative Shares and Equity Judgments</i>	106
<b>Table 2.8</b> Alternative Codings of MRA Predictor Variables, Equity Study	113
<i>Example 2.4 Gender Bias in Faculty Salaries?</i>	116
<b>Table 2.9</b> Mean Faculty Salary at Hypo. U. as f(College, Gender)	117
<b>Table 2.10</b> Data for MRA-Based Anova of Gender-Bias Data Set	118
2.8 Alternatives to the Least-Squares Criterion	121
2.9 Path Analysis	122
2.9.1 Path analytic Terminology	123
2.9.2 Preconditions for Path Analysis	124
2.9.3 Estimating and Testing Path coefficients	126
2.9.4 Decomposition of Correlations into Components	128
2.9.5 Overall Test of Goodness of fit	129
2.9.6 Examples	130
<i>Example 2.5 Mother's Effects on Child's IQ</i>	130
<i>Example 2.6 Gender Bias Revisited: More Light on "Suppression"</i>	134
2.9.7 Some Path-Analysis References	136
Demonstration Problem	136
Answers	139
Some Real Data and a Quiz Thereon	143
<b>Table 2.11</b> Data Set 2: Ratings of Conservatism of Statement	144
Answers	146
<b>Table 2.12</b> Buildup of $R^2$ for Different Orders of Addition of Predictors	146
<b>Figure 2.1</b> Venn diagram of correlations among $Y$ and four predictors	147
Path Analysis Problem	149
Answers to Path Analysis Problem	150

### 3 Hotelling's $T^2$ : Tests on One or Two Mean Vectors

3.1 Single-Sample $t$ and $T^2$	155
<b>Table 3.1</b> Data Set 3: Divisions of Potential Prize, Experiment 3, Harris & Joyce (1980).	157
<i>Example 3.1</i>	162
3.2 Linearly Related Outcome Variables	165
<i>Example 3.2</i>	166
<b>Table 3.2</b> Data Set 4: Results of Deutsch Replication 1	167
3.3 Two-Sample $t$ and $T^2$	170
3.4 Profile Analysis	173
<b>Figure 3.1</b> Response vectors for groups differing in level and slope	174

3.5 Discriminant Analysis	182
3.6 Relationship between $T^2$ and MRA	184
3.7 Assumptions Underlying $T^2$	186
3.7.1 The Assumption of Equal Covariance Matrices	186
3.7.2 Known Covariance Matrix	187
3.7.3 The Assumption of Multivariate Normality	188
3.8 Analyzing Repeated-Measures Designs via $T^2$	188
<b>Table 3.3</b> Repeated-Measures Anova of Data Set 3	189
<i>Example 3.2 Blood Doping</i>	192
<b>Table 3.4</b> 10K Running Time as Affected by an Infusion of One's Own Blood	192
3.9 Single-Symbol Expressions for Simple Cases	196
3.10 Computerized $T^2$	198
3.10.1 Single-Sample and Two-Sample $T^2$	198
3.10.2 Within-Subjects Anova	199
<b>Demonstration Problems</b>	200
<b>Answers</b>	202

## 4 Multivariate Analysis of Variance: Differences Among Several Groups on Several Measures

4.1 One-Way (Univariate) Analysis of Variance	210
4.1.1 The Overall Test	210
<b>Table 4.1</b> Summary Table of Anova on Dependent Variable	213
4.1.2 Specific Comparisons	213
<b>Table 4.2</b> Summary Table for Effects of Instructions on Frequency of <i>DD</i> Outcomes	215
4.2 One-Way Multivariate Analysis of Variance	218
<b>Table 4.3</b> Critical Values for Contrasts Performed on Linear Combinations of Variables	222
4.3 Multiple Profile Analysis	224
<i>Example 4.1 Damselfish Territories</i>	227
<b>Table 4.4</b> Mean Percentage Coverage of Damselfish Territories	227
4.4 Multiple Discriminant Analysis	229
4.5 Greatest Characteristic Roots versus Multiple-Root Tests in Manova	231
4.5.1 "Protected" Univariate Tests	233
4.5.2 Simultaneous Test Procedures and Union Intersection	234
4.5.3 Invalidity of Partitioned- <i>U</i> Tests of Individual Roots	234
4.5.4 Simplified Coefficients as a Solution to the Robustness Problem	237
4.5.5 Finite-Intersection Tests	238

4.6 Simple Cases of Manova	240
4.7 Higher Order Anova: Interactions	243
4.8 Higher Order Manova	245
<i>Example 4.2 Eyeball to Eyeball in a Prisoner's Dilemma</i>	248
<b>Table 4.5</b> Proportion of Mutually Cooperative Choices as f(Contact, Communic'n)	248
<b>Table 4.6</b> Mean Proportion of Total Responses Accounted for by Each Outcome	249
<b>Table 4.7</b> Summary Table for Anova on Discriminant Function from One-Way Manova	251
4.9 Within-Subject Univariate Anova Versus Manova	252
<i>Example 4.3 Stress, Endorphins, and Pain</i>	256
4.10 Computerized Manova	257
4.10.1 Generic Setup for SPSS MANOVA	257
4.10.2 Supplementary Computations	259
4.10.3 Pointing and Clicking to a Manova on SPSS PC	259
4.10.4 Generic Setup for SAS PROC GLM	260
<b>Demonstration Problems</b>	262
Answers	264

## 5 Canonical Correlation: Relationships Between Two Sets of Variables

5.1 Formulae for Computing Canonical Rs	268
5.1.1 Heuristic Justification of Canonical Formulae	270
5.1.2. Simple Cases of Canonical Correlations	272
5.1.3. Example of a Canonical Analysis	274
<b>Table 5.1</b> Correlations of Background Variables with Marijuana Questions	275
<b>Table 5.2</b> Canonical Analysis of Background Variables versus Marijuana Questions	276
5.2 Relationships to Other Statistical Techniques	277
5.3 Likelihood-Ratio Tests of Relationships between Sets of Variables	279
5.4 Generalization and Specialization of Canonical Analysis	280
5.4.1 Testing the Independence of $m$ Sets of Variables	281
<i>Example 5.2 Consistency of Behavior across Different Experimental Games</i>	282
<b>Table 5.3</b> Correlation Matrix for Game Outcome Variables, Flint (1970)	283
5.4.2 Repeated-Battery Canona	284
5.4.3 Rotation of Canonical Variates	288
<i>Example 5.3 A Canonical Cautionary</i>	290
<i>Figure 5.1</i> Naturally Occurring versus Canona-based Pairings of Beefy-Breasted Bowery Birds (BBBs)	291
5.4.4 The Redundancy Coefficient	293
5.4.5 What's Missing from Canonical Analysis?	295

5.5 Computerized Canonical Correlation	297
5.5.1 Matrix-Manipulation Systems	297
5.5.1.1 <i>MATLAB</i>	297
5.5.1.2 <i>SAS PROC MATRIX and SPSS Matrix/End Matrix</i>	299
5.5.2 SAS PROC CANCORR	301
5.5.3. Canona via SPSS MANOVA	304
5.5.4 SPSS Canona From Correlation Matrix : Be Careful	305
<b>Demonstration Problems and Some Real Data Employing Canonical Correlation</b>	<b>307</b>
Answers	309

## 6 Principal Component Analysis: Relationships Within a Single Set of Variables

6.1 Definition of Principal Components	319
6.1.1 Terminology and Notation in PCA and FA	320
6.1.2 Scalar Formulae for Simple Cases of PCA	322
6.1.3 Computerized PCA	325
6.1.4 Additional Unique Properties (AUPs) of PCs	326
6.2 Interpretation of Principal Components	327
<i>Example 6.1 Known generating variables</i>	332
6.3 Uses of Principal Components	333
6.3.1 Uncorrelated Contributions	333
6.3.2 Computational Convenience	334
6.3.3 Principal Component Analysis as a Means of Handling Linear Dependence	335
6.3.4 Examples of PCA	338
<i>Example 6.2 Components of the WISC-R</i>	338
<i>Example 6.3 Attitudes toward cheating</i>	343
<b>Table 6.1</b> PCA on Questions 12-23 of Cheating Questionnaire	344
<i>Example 6.4 Fat, four-eyed, and female again</i>	344
<b>Table 6.2</b> Manova Test of Obesity Main Effect	345
<b>Table 6.3</b> PCA-Based Manova of Obesity Main Effect	347
6.3.5 Quantifying Goodness of Interpretation of Components	348
6.4 Significance Tests for Principal Components	351
6.4.1 Sampling Properties of Covariance-Based PCs	353
6.4.2 Sampling Properties of Correlation-Based PCs	354
6.5 Rotation of Principal Components	356
<i>Example 6.1 revisited</i>	356
6.5.1 Basic Formulae for Rotation	358
<b>Figure 6.1</b> Factor structures, example 6.1	358
<b>Figure 6.2</b> Rotation, general case	358
6.5.2 Objective Criteria for Rotation	360

<b>Table 6.4</b> Quadrant within which $4\phi$ Must Fall as Function of Signs of Numerator and Denominator of Expression (6.9)	364
6.5.3 Examples of Rotated PCs	365
<b>Table 6.5</b> Intermediate Calculations for Quartimax and Varimax Rotation	365
<b>Table 6.6</b> Varimax Rotation of $PC_1$ - $PC_4$ , Cheating Questionnaire	367
<b>Table 6.7</b> Varimax Rotation of All Twelve PCs, Cheating Questionnaire	368
<b>Table 6.8</b> Large Loadings for Cheating Questionnaire	369
6.5.4 Individual Scores on Rotated PCs	369
<b>Example 6.5</b> A factor fable	373
<b>Figure 6.3</b> Architectural dimensions of houses	373
<b>Figure 6.4</b> Schematic representation of 27 houses	374
<b>Table 6.9</b> Scores on Observed and Derived Variables for 27 Houses	375
<b>Figure 6.5</b> Same 27 houses sorted on basis of loadings-based interpretation of Factor 1	377
6.5.5 Uncorrelated-Components Versus Orthogonal-Profiles Rotation	379
<b>Demonstration Problems</b>	381
Answers	383
<b>Figure 6.6</b> Rotation of factor structure for problem 1	386

## 7 Factor Analysis: The Search for Structure

7.1 The Model	394
7.2 Communalities	397
7.2.1 Theoretical Solution	398
7.2.2 Empirical Approximations	400
7.2.3 Iterative Procedure	401
7.2.4 Is the Squared Multiple Correlation the True Communality?	401
7.3 Factor Analysis Procedures Requiring Communality Estimates	404
7.3.1 Principal Factor Analysis	404
7.3.2 Triangular (Choleski) Decomposition	405
7.3.3 Centroid Analysis	406
7.4 Methods Requiring Estimate of Number of Factors	406
7.5 Other Approaches to Factor Analysis	409
7.6 Factor Loadings versus Factor Scores	410
7.6.1 Factor Score Indeterminacy	411
7.6.2 Relative Validities of Loadings-Derived versus Scoring-Coefficient-Derived Factor Interpretations	412
<b>Table 7.1</b> Mean Validity, Univocality, and Orthogonality of Regression and Loading Estimates for Three Levels of Complexity	413
7.6.3 Regression-Based Interpretation of Factors is Still a Hard Sell	414
7.7 Relative Merits of Principal Component Analysis versus Factor Analysis	416
7.7.1 Similarity of Factor Scoring Coefficients	416
<b>Table 7.2</b> Comparison of Factor Structures for PCA versus Two PFAs of Same Data	417

<b>Table 7.3</b> Comparison of Kaiser-Normalized Factor Structures	418
<b>Table 7.4</b> Comparison of Factor-Score Coefficients	418
7.7.2 Bias in Estimates of Factor Loadings	420
7.8 Computerized Exploratory Factor Analysis	421
<i>Example 7.1 WISC-R Revisited</i>	423
7.9 Confirmatory Factor Analysis	433
7.9.1 SAS PROC CALIS	
<i>Example 7.1 Revisited: Model Comparisons Galore</i>	434
<b>8 The Forest Revisited</b>	
8.1 Scales of Measurement and Multivariate Statistics	444
<b>Table 8.1</b> Representative Critical Values for Measures of Association	446
8.2 Effects of Violations of Distributional Assumptions in Multivariate Analysis	450
8.3 Nonlinear Relationships in Multivariate Statistics	453
8.4 The Multivariate General Linear Hypothesis	456
<i>Example 8.1 Unbalanced Manova via the multivariate general linear model</i>	460
8.5 Structural Equation Modeling	464
8.5.1 General Approach and Examples	464
<i>Example 8.2 Path Analysis of Scarr (1985) via SEM</i>	464
<i>Example 8.3 All Three Colleges in the Faculty Salary Example</i>	468
<i>Example 8.4 Increment to Canonical <math>R^2</math> via CALIS LinEqs?</i>	470
8.5.2 SEM Is Not a General Model for Multivariate Statistics	473
<i>Example 8.5 Higher-Order Confirmatory Factor Analysis via SEM: WISC-R One More Time</i>	473
8.5.3 Other User-Friendly SEM Programs	478
8.6 Where to Go from Here	479
8.7 Summing Up	480

## Digression 1

### Finding Maxima and Minima of Polynomials

D1.1 Derivatives and Slopes	482
D1.2 Optimization Subject to Constraints	485

## Digression 2

### Matrix Algebra

D2.1 Basic Notation	487
D2.2 Linear Combinations of Matrices	489
D2.3 Multiplication of Matrices	489



D2.4 Permissible Manipulations	493
D2.5 Inverses	493
D2.6 Determinants	496
D2.7 Some Handy Formulae for Inverses and Determinants in Simple Cases	500
D2.8 Rank	501
D2.9 Matrix Calculus	502
D2.10 Partitioned Matrices	503
D2.11 Characteristic Roots and Vectors	506
D2.12 Solution of Homogeneous Systems of Equations	512

## Digression 3

### Solution of Cubic Equations 514

## Appendix A

### Statistical Tables

A.1 - A.4 (Why omitted from this edition)	517
A.5 Greatest Characteristic Root Distribution	518

## Appendix B

### Computer Programs Available from the Author

B.1 <i>cvinter</i> : $p$ values and Critical Values for Univariate Statistics	532
B.2 <i>gcrinter</i> : Critical Values for the Greatest Characteristic Root (g.c.r.) Distribution	532

## Appendix C

### Derivations

Derivation 1.1 Per-Experiment and Experimentwise Error Rates for Bonferroni-Adjusted Tests	533
Derivation 2.1 Scalar Formulae for MRA with One, Two, and Three Predictors	536
Derivation 2.2 Coefficients That Minimize Error Also Maximize Correlation	539
Derivation 2.3 Maximizing $r$ via Matrix Algebra	541
Derivation 2.4 Independence of Irrelevant Parameters	542
Derivation 2.5 Variances of $b_j$ s and of Linear Combinations Thereof	542
Derivation 2.6 Drop in $R^2 = b_j^2(1 - R_{j, \text{oth}}^2)$	543
Derivation 2.7 MRA on Group-Membership Variables Yields Same $F$ As Anova	544
Derivation 2.8 Unweighted Means and Least-Squares Anova Are Identical in the 2 <sup>n</sup> Design	545

Derivation 3.1 $T^2$ and Associated Discriminant Function	546
Single-Sample $T^2$	546
Two-Sample $T^2$	548
Derivation 3.2 Relationship between $T^2$ and MRA	549
Two-Sample $t$ Versus Pearson $r$ With Group-Membership Variables	549
Single-Sample $t$ Test versus "Raw-Score" $r_{cv}$	550
$T^2$ Versus MRA	551
Derivation 4.1 Maximizing $F(\mathbf{a})$ in Manova	552
Derivation 5.1 Canonical Correlation and Canonical Variates	554
Derivation 5.2 Canonical Correlation as "Mutual Regression Analysis"	556
Derivation 5.3 Relationship between Canonical Analysis and Manova	557
Derivation 6.1 Principal Components	560
Derivation 6.2 PC Coefficients Define Both Components in Terms of Xs and Xs in Terms of PCs	562
Derivation 6.3 What Does Rotation of Loadings Do to Coefficients?	564
Derivation 7.1 Near Equivalence of PCA and Equal-Communalities PFA	566

## **References 567**

## **Index 584**