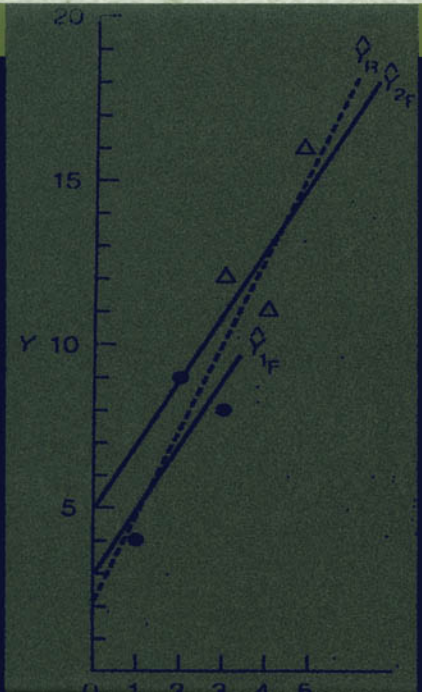
CD Enclosed

Designing Experiments and Analyzing Data

Second Edition

A Model Comparison Perspective



Scott E. Maxwell
Harold D. Delaney

TLFeBOOK

Contents

Preface xvii

I CONCEPTUAL BASES OF EXPERIMENTAL DESIGN AND ANALYSIS

1 The Logic of Experimental Design 3
The Traditional View of Science 3
Responses to the Criticisms of the Idea of Pure Science 5
Assumptions 5
Modern Philosophy of Science 10
Threats to the Validity of Inferences from Experiments 22
Types of Validity 23
Conceptualizing and Controlling for Threats to Validity 30
Exercises 32

2 Introduction to the Fisher Tradition 34
"Interpretation and its Reasoned Basis" 35
A Discrete Probability Example 37
Randomization Test 41
Of Hypotheses and p Values: Fisher Versus Neyman–Pearson 47
Toward Tests Based on Distributional Assumptions 49
Statistical Tests with Convenience Samples 49
The Assumption of Normality 50
Overview of Experimental Designs to be Considered 56
Exercises 59

II MODEL COMPARISONS FOR BETWEEN-SUBJECTS DESIGNS

3 Introduction to Model Comparisons: One-Way Between-Subjects Designs 67
The General Linear Model 69
One-Group Situation 71
Basics of Models 71
Proof That $\bar{Y}$ Is the Least-Squares Estimate of μ (*Optional*) 73
Development of the General Form of the Test Statistic 75
Numerical Example 78
Relationship of Models and Hypotheses 80
Two-Group Situation 80
Development in Terms of Models 80

TLFeBOOK

Alternative Development and Identification with Traditional Terminology 83
Tests of Replication (*Optional*) 85
The General Case of One-Way Designs 88
Formulation in Terms of Models 88
Numerical Example 91
A Model in Terms of Effects 93
On Tests of Significance and Measures of Effect 98
Measures of Effect 100
Measures of Effect Size 101
Measures of Association Strength 104
Alternative Representations of Effects 107
Statistical Assumptions 110
Implications for Expected Values 110
Robustness of ANOVA 111
Checking for Normality and Homogeneity of Variance 114
Transformations 117
Power of the *F* Test: One-Way ANOVA 120
Determining Sample Size Using *d* and Table 3.10 123
Pilot Data and Observed Power 124
Exercises 126
Extension: Robust Methods for One-Way Between-Subject Designs 129
Parametric Modifications 131
Nonparametric Approaches 136
Choosing Between Parametric and Nonparametric Tests 137
Two Other Approaches (*Optional*) 143
Why Does the Usual *F* Test Falter with Unequal *n*s When Population Variances Are Unequal? (*Optional*) 145
Exercises 147

4 Individual Comparisons of Means 149
A Model Comparison Approach for Testing Individual Comparisons 150
Preview of Individual Comparisons 150
Relationship to Model Comparisons 150
Derivation of Parameter Estimates and Sum of Squared Errors (*Optional*) 152
Expression of *F* Statistic 153
Numerical Example 155
Complex Comparisons 157
Models Perspective 157
Numerical Example 162
The *t* Test Formulation of Hypothesis Testing for Contrasts 163
Practical Implications 164
Unequal Population Variances 165
Numerical Example 168
Measures of Effect 169
Measures of Effect Size 170
Measures of Association Strength 173
Testing More Than One Contrast 177
How Many Contrasts Should Be Tested? 177
Linear Independence of Contrasts 178
Orthogonality of Contrasts 179

TLFeBOOK

Example of Correlation Between Nonorthogonal Contrasts (*Optional*) 180
Another Look at Nonorthogonal Contrasts: Venn Diagrams 182
Exercises 186
Extension: Derivation of Sum of Squares for a Contrast 190

5 Testing Several Contrasts: The Multiple-Comparison Problem 193
Multiple Comparisons 193
Experimentwise and Per-Comparison Error Rates 193
Simultaneous Confidence Intervals 196
Levels of Strength of Inference 197
Types of Contrasts 198
Overview of Techniques 199
Planned Versus Post Hoc Contrasts 200
Multiple Planned Comparisons 201
Bonferroni Adjustment 202
Modification of the Bonferroni Approach With Unequal Variances 205
Numerical Example 206
Pairwise Comparisons 208
Tukey's WSD Procedure 210
Modifications of Tukey's WSD 212
Numerical Example 213
Post Hoc Complex Comparisons 213
Proof That $SS_{max} = SS_B$ 215
Comparison of Scheffé to Bonferroni and Tukey 217
Modifications of Scheffé's Method 218
Numerical Example 219
Other Multiple-Comparison Procedures 221
Dunnett's Procedure for Comparisons with a Control 221
Numerical Example 222
Procedures for Comparisons with the Best 223
Numerical Example 227
Fisher's LSD (Protected *t*) 229
False Discovery Rate 230
Choosing an Appropriate Procedure 234
Exercises 237

6 Trend Analysis 243
Quantitative Factors 243
Statistical Treatment of Trend Analysis 244
The Slope Parameter 246
Numerical Example 248
Hypothesis Test of Slope Parameter 249
Confidence Interval and Other Effect Size Measures for Slope Parameter 251
Numerical Example 251
Testing for Nonlinearity 254
Numerical Example 257
Testing Individual Higher Order Trends 257
Contrast Coefficients for Higher Order Trends 259
Numerical Example 260

TLFeBOOK

Further Examination of Nonlinear Trends 263
Trend Analysis with Unequal Sample Sizes 267
Concluding Comments 269
Exercises 269

7 Two-Way Between-Subjects Factorial Designs 275
The 2 × 2 Design 275
The Concept of Interaction 277
Additional Perspectives on the Interaction 278
A Model Comparison Approach to the General Two-Factor Design 280
Alternate Form of Full Model 280
Comparison of Models for Hypothesis Testing 284
Numerical Example 290
Familywise Control of Alpha Level 291
Measures of Effect 291
Follow-Up Tests 297
Further Investigation of Main Effects 297
Marginal Mean Comparisons Without Homogeneity Assumption (*Optional*) 300
Further Investigation of an Interaction—Simple Effects 301
An Alternative Method for Investigating an Interaction—Interaction Contrasts 309
Statistical Power 317
Advantages of Factorial Designs 319
Nonorthogonal Designs 320
Design Considerations 321
Relationship Between Design and Analysis 321
Analysis of the 2 × 2 Nonorthogonal Design 322
Test of the Interaction 322
Unweighted Marginal Means and Type III Sum of Squares 324
Unweighted Versus Weighted Marginal Means 325
Type II Sum of Squares 327
Summary of Three Types of Sum of Squares 328
Analysis of the General $a \times b$ Nonorthogonal Design 329
Test of the Interaction 329
Test of Unweighted Marginal Means 330
Test of Marginal Means in an Additive Model 331
Test of Weighted Marginal Means 332
Summary of Types of Sum of Squares 333
Which Type of Sum of Squares Is Best? 334
A Note on Statistical Packages for Analyzing Nonorthogonal Designs 335
Numerical Example 337
Final Remarks 343
Exercises 343

8 Higher Order Between-Subjects Factorial Designs 354
The 2 × 2 × 2 Design 354
The Meaning of Main Effects 355
The Meaning of Two-Way Interactions 356

TLFeBOOK

The Meaning of the Three-Way Interaction 357
Graphical Depiction 359
Further Consideration of the Three-Way Interaction 361
Summary of Meaning of Effects 366
The General $A \times B \times C$ Design 367
The Full Model 367
Formulation of Restricted Models 368
Numerical Example 372
Implications of a Three-Way Interaction 374
General Guideline for Analyzing Effects 376
Summary of Results 381
Graphical Depiction of Data 382
Confidence Intervals for Single Degree of Freedom Effects 383
Other Questions of Potential Interest 386
Tests to Be Performed When the Three-Way Interaction Is Nonsignificant 387
Nonorthogonal Designs 389
Higher Order Designs 391
Exercises 392

9 Designs With Covariates: ANCOVA and Blocking 399
ANCOVA 401
The Logic of ANCOVA 401
Linear Models for ANCOVA 403
Two Consequences of Using ANCOVA 414
Assumptions in ANCOVA 420
Numerical Example 428
Measures of Effect 431
Comparisons Among Adjusted Group Means 434
Generalizations of the ANCOVA Model 438
Choosing Covariates in Randomized Designs 439
Sample Size Planning and Power Analysis in ANCOVA 441
Alternate Methods of Analyzing Designs with Concomitant Variables 443
ANOVA of Residuals 444
Gain Scores 444
Blocking 448
Exercises 453
Extension: Heterogeneity of Regression 456
Test for Heterogeneity of Regression 456
Accommodating Heterogeneity of Regression 460

10 Designs with Random or Nested Factors 469
Designs with Random Factors 469
Introduction to Random Effects 469
One-Factor Case 471
Two-Factor Case 474
Numerical Example 481
Alternative Tests and Design Considerations with Random Factors 483
Follow-up Tests and Confidence Intervals 484

TLFeBOOK

Measures of Association Strength 485
Using Statistical Computer Programs to Analyze Designs with Random Factors 489
Determining Power in Designs with Random Factors 490
Designs with Nested Factors 494
Introduction to Nested Factors 494
Example 499
Models and Tests 499
Degrees of Freedom 504
Statistical Assumptions and Related Issues 506
Follow-up Tests and Confidence Intervals 507
Strength of Association in Nested Designs 508
Using Statistical Computer Programs to Analyze Nested Designs 509
Selection of Error Terms When Nested Factors Are Present 510
Complications That Arise in More Complex Designs 512
Exercises 517

III MODEL COMPARISONS FOR DESIGNS INVOLVING WITHIN-SUBJECTS FACTORS

11 One-Way Within-Subjects Designs: Univariate Approach 525
Prototypical Within-Subjects Designs 525
Advantages of Within-Subjects Designs 527
Analysis of Repeated Measures Designs with Two Levels 527
The Problem of Correlated Errors 527
Reformulation of Model 529
Analysis of Within-Subjects Designs with More Than Two Levels 531
Traditional Univariate (Mixed-Model) Approach 532
Comparison of Full and Restricted Models 533
Estimation of Parameters: Numerical Example 534
Assumptions in the Traditional Univariate (Mixed-Model) Approach 539
Homogeneity, Sphericity, and Compound Symmetry 540
Numerical Example 541
Adjusted Univariate Tests 542
Lower-Bound Adjustment 543
$\hat{\varepsilon}$ Adjustment 543
$\tilde{\varepsilon}$ Adjustment 544
Summary of Four Mixed-Model Approaches 545
Measures of Effect 547
Comparisons Among Individual Means 550
Confidence Intervals for Comparisons 551
Confidence Intervals with Pooled and Separate Variances (*Optional*) 553
Considerations in Designing Within-Subjects Experiments 555
Order Effects 556
Differential Carryover Effects 556
Controlling for Order Effects with More Than Two Levels: Latin Square Designs 557

TLFeBOOK

Relative Advantages of Between-Subjects and Within-Subjects Designs 561
Intraclass Correlations for Assessing Reliability 563
Exercises 567

12 Higher-Order Designs with Within-Subjects Factors: Univariate Approach 573
Designs with Two Within-Subjects Factors 573
Omnibus Tests 574
Numerical Example 577
Further Investigation of Main Effects 579
Further Investigation of an Interaction—Simple Effects 581
Interaction Contrasts 582
Statistical Packages and Pooled Error Terms Versus Separate Error Terms 583
Assumptions 584
Adjusted Univariate Tests 588
Confidence Intervals 590
Quasi-F Ratios 590
One Within-Subjects Factor and One Between-Subjects Factor in the Same Design 592
Omnibus Tests 593
Further Investigation of Main Effects 599
Further Investigation of an Interaction—Simple Effects 601
Interaction Contrasts 605
Assumptions 607
Adjusted Univariate Tests 609
More Complex Designs 610
Designs with Additional Factors 610
Latin Square Designs 611
Exercises 616

13 One-Way Within-Subjects Designs: Multivariate Approach 624
A Brief Review of Analysis for Designs with Two Levels 624
Multivariate Analysis of Within-Subjects Designs with Three Levels 626
Need for Multiple D Variables 626
Full and Restricted Models 627
The Relationship Between D_1 and D_2 629
Matrix Formulation and Determinants 630
Test Statistic 632
Multivariate Analysis of Within-Subjects Designs with a Levels 633
Forming D Variables 633
Test Statistic 635
Numerical Example 635
Measures of Effect 638
Choosing an Appropriate Sample Size 639
Choice of D Variables 645
Tests of Individual Contrasts 647
Quantitative Repeated Factors (*Optional*) 649

TLFeBOOK

Multiple-Comparison Procedures: Determination of Critical Values 650
Planned Comparisons 651
Pairwise Comparisons 651
Post Hoc Complex Comparisons 652
Finding D_{max} (*Optional*) 653
Confidence Intervals for Contrasts 654
The Relationship Between the Multivariate Approach and the Mixed-Model Approach 658
Orthonormal Contrasts 658
Comparison of the Two Approaches 660
Reconceptualization of ε in Terms of $E^*(F)$ (*Optional*) 663
Multivariate and Mixed-Model Approaches for Testing Contrasts 665
Numerical Example 666
The Difference in Error Terms 667
Which Error Term Is Better? 669
A General Comparison of the Multivariate and Mixed-Model Approaches 671
Assumptions 672
Tests of Contrasts 672
Type I Error Rates 672
Type II Error Rates 673
Summary 675
Exercises 676

14 Higher Order Designs with Within-Subjects Factors: Multivariate Approach 682
Two Within-Subjects Factors, Each with Two Levels 682
Formation of Main-Effect D Variables 684
Formation of Interaction D Variables 686
Relationship to the Mixed-Model Approach 688
Multivariate Analysis of Two-Way $a \times b$ Within-Subjects Designs 688
Formation of Main-Effect D Variables 688
Formation of Interaction D Variables 691
Omnibus Tests—Multivariate Significance Tests 693
Measures of Effect 694
Further Investigation of Main Effects 695
Further Investigation of an Interaction—Simple Effects 696
Interaction Contrasts 698
Confidence Intervals for Contrasts 699
The Relationship Between the Multivariate and the Mixed-Model Approaches (*Optional*) 701
Multivariate and Mixed-Model Approaches for Testing Contrasts 703
Comparison of the Multivariate and Mixed-Model Approaches 704
One Within-Subjects Factor and One Between-Subjects Factor in the Same Design 704
Split-Plot Design With Two Levels of the Repeated Factor 704
General $a \times b$ Split-Plot Design 713
Measures of Effect 725

TLFeBOOK

Confidence Intervals for Contrasts 738
The Relationship Between the Multivariate and the Mixed-Model Approaches (*Optional*) 742
Assumptions of the Multivariate Approach 744
Multivariate and Mixed-Model Approaches for Testing Within-Subjects Contrasts 745
Comparison of the Multivariate and Mixed-Model Approaches 746
More Complex Designs (*Optional*) 746
Exercises 752

IV ALTERNATIVE ANALYSIS STRATEGIES

15 An Introduction to Multilevel Models for Within-Subjects Designs 763
Advantages of New Methods 763
Within-Subjects Designs 763
Overview of Remainder of Chapter 764
Within-Subjects Designs 765
Various Types of Within-Subjects Designs 765
Models for Longitudinal Data 765
Review of the ANOVA Mixed-Model Approach 766
Random Effects Models 767
Maximum Likelihood Approach 767
An Example of Maximum Likelihood Estimation (*Optional*) 768
Comparison of ANOVA and Maximum Likelihood Models 770
Numerical Example 773
A Closer Look at the Random Effects Model 778
Graphical Representation of Longitudinal Data 779
Graphical Representation of the Random Intercept Model 781
Coding Random Effects Predictor Variables 785
Random Effects Parameters 786
Numerical Example 788
Graphical Representation of a Model With Random Slope and Intercept 790
Further Consideration of Competing Models 791
Additional Models Deserving Consideration 793
Graphical Representation of a Growth Curve Model 798
Design Considerations 800
An Alternative to the Random Effects Model 802
Additional Covariance Matrix Structures 809
Tests of Contrasts 813
Overview of Broader Model Comparison 814
Complex Designs 816
Factorial Fixed Effects 817
Multiple Variables Measured Over Time 818
Unbalanced Designs 818
Conclusion 820
Exercises 820

TLFeBOOK

16 An Introduction to Multilevel Hierarchical Mixed Models: Nested Designs **828**
Review of the ANOVA Approach **829**
Maximum Likelihood Analysis Models for the Simple Nested Design **831**
Numerical Example—Equal *n* **833**
Numerical Example—Unequal *n* **840**
Maximum Likelihood Analysis Models for Complex Nested Designs **845**
Hierarchical Representation of the Model for a Simple Nested Design **846**
Models With Additional Level 2 Variables **849**
Models with Additional Level 1 Variables **853**
Exercises **867**

Appendixes

A Statistical Tables **A-1**

B Part 1. Linear Models: The Relation Between ANOVA and Regression **B-1**

Part 2. A Brief Primer of Principles of Formulating and Comparing Models **B-26**

C Notes **C-1**

D Solutions to Selected Exercises **D-1**

E References **E-1**

Name Index **N-1**

Subject Index **S-1**

TLFeBOOK