Contents

1 SCIENCE AND STATISTICS 1

1.1 The Learning Process 1
1.2 The Role of Experimental Design 4
1.3 Difficulties Mitigated by Statistical Methods 7
1.4 A Typical Investigation 9
1.5 How to Use Statistical Techniques 14

PART I COMPARING TWO TREATMENTS

2 USE OF EXTERNAL REFERENCE DISTRIBUTION TO COMPARE TWO MEANS 21

2.1 Relevant Reference Sets and Distributions 21
2.2 Theory: Probability Distributions, Parameters, and Statistics 24
2.3 The Industrial Experiment: External Reference Distribution 31
2.4 Theory: Normal and *t* Distributions 38
2.5 The Industrial Experiment: An External Reference Distribution Based on the *t* Distribution 51

Appendix 2A Calculation of the Sample Average, Sample Variance, and Sample Standard Deviation by Coding Data 53

3 RANDOM SAMPLING AND THE DECLARATION OF INDEPENDENCE 57

3.1 Theory: Statistical Dependence and Independence and the Random Sampling Model 57

3.2 The Industrial Experiment: Reference Distribution Based on Random Sampling Model, External Value for σ 74
3.3 The Industrial Experiment: Reference Distribution Based on Random Sampling Model, Internal Estimate of σ 76
3.4 Summary: What Have We Learned from the Industrial Experiment Example? 82
Appendix 3A Mean and Variance of a Linear Combination of Observations 87
Appendix 3B Robustness of Some Statistical Procedures 89
Appendix 3C Fisher's Concept of Sufficiency 91

4 RANDOMIZATION AND BLOCKING WITH PAIRED COMPARISONS 93

4.1 Randomization to the Rescue: Tomato Plant Example 93
4.2 Randomized Paired Comparison Design: Boys' Shoes Example 97
4.3 Blocking and Randomization 102
4.4 Noise Structure, Models, and Randomization 104
4.5 Summary: Comparison, Replication, Randomization, and Blocking in Simple Comparative Experiments 105

5 SIGNIFICANCE TESTS AND CONFIDENCE INTERVALS FOR MEANS, VARIANCES, PROPORTIONS, AND FREQUENCIES 107

5.1 A More Detailed Discussion of Significance Tests 107
5.2 Confidence Intervals for a Difference in Means: Paired Comparison Design 110
5.3 Confidence Intervals for a Difference in Means: Unpaired Design 115
5.4 Inferences about Variances of Normally Distributed Data 117
5.5 Inferences about Proportions: The Binomial Distribution 123
5.6 Inferences about Frequencies: The Poisson Distribution 137
5.7 Contingency Tables and Tests of Association 145

PROBLEMS FOR PART I 152

PART II COMPARING MORE THAN TWO TREATMENTS

6 EXPERIMENTS TO COMPARE k TREATMENT MEANS 165

6.1 Blood Coagulation Times with Four Different Diets 165
6.2 Estimating the Amount of Variation Within and Between Treatments 167
6.3 The Arithmetic and Geometry of the Analysis of Variance Table 170
6.4 Decomposition of the Observations Implied by the Analysis 175
6.5 Diagnostic Checking of the Basic Model 182
6.6 Use of the Analysis of Variance Table 187
6.7 Use of a Reference Distribution to Compare Means 190
6.8 Summary 193
Appendix 6A Shortcut Method for Constructing the Analysis of Variance Table 194
Appendix 6B Vectors and Geometry Associated with the Analysis of a Sample 197
Appendix 6C Multiple Comparisons 203

7 RANDOMIZED BLOCKS AND TWO-WAY FACTORIAL DESIGNS 208

7.1 Example: Comparison of Four Variants of a Penicillin Production Process 209
7.2 A Model with Corresponding Decomposition of Observations 210
7.3 Implications of the Additive Model 218
7.4 Diagnostic Checking of the Model 220
7.5 Use of the Analysis of Variance Table 223
7.6 The Use of Reference Distributions To Compare Individual Means 226
7.7 A Two-Way (Factorial) Design 228
7.8 Simplification and Increased Sensitivity from Transformation 231
7.9 Likelihood Estimation of the Transformation 239
7.10 Summary 241
Appendix 7A Calculations for Constructing Analysis of Variance Table for Randomized Block Design 241

Appendix 7B Algebraic Demonstration of the Additivity of the Sums of Squares in a Randomized Block 243

8 DESIGNS WITH MORE THAN ONE BLOCKING VARIABLE 245

8.1 Latin Square Designs: Automobile Emissions and Synthetic Yarn Examples 245
8.2 Graeco- and Hyper-Graeco–Latin Squares: First Wear Testing Example 255
8.3 Balanced Incomplete Block Designs: Second Wear Testing Example 258
Appendix 8A Some Useful Latin Squares and How to Use Them to Construct Graeco–Latin and Hyper-Graeco–Latin Square Design 261
Appendix 8B Analysis of Variance for $k \times k$ Latin Square Designs with r Replicates 263
Appendix 8C Some Useful Balanced Incomplete Block Designs 269
Appendix 8D Analysis of Variance and Computation of Adjusted Treatment Averages for Balanced Incomplete Block Designs 275

PROBLEMS FOR PART II 281

PART III MEASURING THE EFFECTS OF VARIABLES

9 EMPIRICAL MODELING 291

9.1 Mathematical Models 291
9.2 Geometric Representation of Empirical Relationships 296
9.3 The Problem of Experimental Design 298
9.4 Comprehensive Versus Sequential Approach to Experimental Investigations 303

10 FACTORIAL DESIGNS AT TWO LEVELS 306

10.1 General Factorial Designs and Designs at Two Levels 306
10.2 An Example of a 2^3 Factorial Design: Pilot Plant Investigation 307

10.3 Calculation of Main Effects 309
10.4 Interaction Effects 313
10.5 Interpretation of Results 317
10.6 Calculation of Standard Errors for Effects Using Replicated Runs 319
10.7 Quicker Methods for Calculating Effects 322
10.8 A 2^4 Factorial Design: Process Development Study 324
10.9 Analysis of Factorials Using Normal Probability Paper 329
10.10 Transformation of Data from Factorial Designs 334
10.11 Blocking 336
10.12 Summary 342
Appendix 10A Yates's Algorithm 342
Appendix 10B More on Blocking Factorial Designs 344

11 MORE APPLICATIONS OF FACTORIAL DESIGNS 352

11.1 Example 1: The Effects of Three Variables on Clarity of Film 352
11.2 Example 2: The Effects of Three Variables on Physical Properties of a Polymer Solution 353
11.3 Example 3: Development of Screening Facility for Storm Water Overflows 354
11.4 Example 4: Simple Factorials Used Sequentially in Evolutionary Operation—Petrochemical Plant 362
11.5 Example 5: Simple Factorials Used Sequentially in Evolutionary Operation—Polymer Unit 365
11.6 Summary 368
Appendix 11A A Suggested Exercise 368

12 FRACTIONAL FACTORIAL DESIGNS AT TWO LEVELS 374

12.1 Redundancy 374
12.2 A Half-Fraction of a 2^5 Design: Reactor Example 376
12.3 Construction and Analysis of Half-Fractions: Reactor Example 381
12.4 The Concept of Design Resolution: Reactor Example 385
12.5 Resolution III Designs: Bicycle Example 390
12.6 Resolution IV Designs: Injection Molding Example 398
12.7 Elimination of Block Effects in Fractional Designs 404

12.8 Designs of Resolution V and Higher 407
12.9 Summary 409
Appendix 12A Structure of the Fractional Designs 409
Appendix 12B Choosing Additional Runs To Resolve Ambiguities from Fractional Factorials 413

13 MORE APPLICATIONS OF FRACTIONAL FACTORIAL DESIGNS 419

13.1 Example 1: Effects of Five Variables on Some Properties of Cast Films 419
13.2 Example 2: Stability of New Product 422
13.3 Example 3: Bottleneck at the Filtration Stage of an Industrial Plant 424
13.4 Example 4: Sensitivity Analysis of a Simulation Model—Controller–Aircraft System 429
13.5 Summary 432

PROBLEMS FOR PART III 434

PART IV BUILDING MODELS AND USING THEM

14 SIMPLE MODELING WITH LEAST SQUARES (REGRESSION ANALYSIS) 453

14.1 One-Parameter Model (Straight Line through the Origin): Aerosol Example 453
14.2 Two-Parameter Model: Impurity Example 462
14.3 Straight Line Model: Welding Example 473
14.4 General Case for Models Linear in the Parameters 479
14.5 Polynomial Model: Growth Rate Example 480
14.6 Nonlinear Model: Biochemical Oxygen Demand Example 483
14.7 Hazards of Fitting Regression Equations to Happenstance Data 487
Appendix 14A Why Do the Normal Equations Yield Least Squares Estimates? 498
Appendix 14B Matrix Version of the Normal Equations 501

Appendix 14C Analysis of Factorials, Botched and Otherwise 503
Appendix 14D Unweighted and Weighted Least Squares 505

15 RESPONSE SURFACE METHODS 510

15.1 Weakness of Classical One-Variable-at-a-Time Strategy: Chemical Example 510
15.2 Illustration of Response Surface Methodology: Chemical Example 513
15.3 A Specification Problem 526
15.4 Maxima, Ridges, and Canonical Analysis 526
15.5 Applications of Response Surface Methods 534
15.6 Summary 535

16 MECHANISTIC MODEL BUILDING 540

16.1 Empirical and Mechanistic Models 540
16.2 Possible Advantages of Mechanistic Models 544
16.3 Techniques for Mechanistic Modeling 546
16.4 The Model-Building Process 548
16.5 Model Testing with Diagnostic Parameters 550
16.6 Importance of Plotting Data in the Age of Computers 552
16.7 Summary 552

17 STUDY OF VARIATION 556

17.1 Graphs and Control Charts: Impurity Determination Example 556
17.2 Transmission of Error 563
17.3 Variance Components: Pigment Paste Example 571
Appendix 17A Calculating Variance Components from an Analysis of Variance Table 581

18 MODELING DEPENDENCE: TIME SERIES 584

18.1 The Industrial Data of Chapter 2 Reconsidered as a Time Series 585
18.2 Statistical Modeling Revisited 588

18.3 Forecasting: Refrigerator Sales Example 591
18.4 Feedback Control: Dye Level Example 598
18.5 Intervention Analysis: Los Angeles Air Pollution Example 602
Appendix 18A Derivation of Equation 18.4 604

PROBLEMS FOR PART IV 606

APPENDIX: TABLES 629

INDEX 645