Statistics for Experimenters
Design, Innovation, and Discovery

Second Edition

GEORGE E. P. BOX
J. STUART HUNTER
WILLIAM G. HUNTER
## Contents

**Preface to the Second Edition** xv

**Chapter 1 Catalyzing the Generation of Knowledge** 1

1.1. The Learning Process 1
1.2. Important Considerations 5
1.3. The Experimenter's Problem and Statistical Methods 6
1.4. A Typical Investigation 9
1.5. How to Use Statistical Techniques 13
References and Further Reading 14

**Chapter 2 Basics (Probability, Parameters, and Statistics)** 17

2.1. Experimental Error 17
2.2. Distributions 18
2.3. Statistics and Parameters 23
2.4. Measures of Location and Spread 24
2.5. The Normal Distribution 27
2.6. Normal Probability Plots 33
2.7. Randomness and Random Variables 34
2.8. Covariance and Correlation as Measures of Linear Dependence 37
2.9. Student's t Distribution 39
2.10. Estimates of Parameters 43
2.11. Random Sampling from a Normal Population 44
2.12. The Chi-Square and F Distributions 46
2.13. The Binomial Distribution 48
2.14. The Poisson Distribution 54

ix
5.2. Example 1: The Effects of Three Factors (Variables) on Clarity of Film 174
5.3. Example 2: The Effects of Three Factors on Three Physical Properties of a Polymer Solution 175
5.4. A $2^3$ Factorial Design: Pilot Plant Investigation 177
5.5. Calculation of Main Effects 178
5.6. Interaction Effects 181
5.7. Genuine Replicate Runs 183
5.8. Interpretation of Results 185
5.9. The Table of Contrasts 186
5.10. Misuse of the ANOVA for $2^k$ Factorial Experiments 188
5.11. Eyeing the Data 190
5.12. Dealing with More Than One Response: A Pet Food Experiment 193
5.15. Other Models for Factorial Data 208
5.16. Blocking the $2^k$ Factorial Designs 211
5.17. Learning by Doing 215
5.18. Summary 219

Appendix 5A. Blocking Larger Factorial Designs 219
Appendix 5B. Partial Confounding 221
References and Further Reading 222

Chapter 6 Fractional Factorial Designs 235

6.1. Effects of Five Factors on Six Properties of Films in Eight Runs 235
6.2. Stability of New Product, Four Factors in Eight Runs, a $2^4-1$ Design 236
6.3. A Half-Fraction Example: The Modification of a Bearing 239
6.4. The Anatomy of the Half Fraction 240
6.5. The $2^7-4$ Design: A Bicycle Example 244
6.6. Eight-Run Designs 246
6.7. Using Table 6.6: An Illustration 247
6.8. Sign Switching, Foldover, and Sequential Assembly 249
6.9. An Investigation Using Multiple-Column Foldover 252
6.10. Increasing Design Resolution from III to IV by Foldover 257
6.11. Sixteen-Run Designs 258
6.12. The $2^{5-1}$ Nodal Half Replicate of the $2^5$ Factorial: Reactor Example 259
6.13. The $2^{8-4}$ Nodal Sixteenth Fraction of a $2^8$ Factorial 263
6.14. The $2^{15-11}$ Nodal Design: The Sixty-Fourth Fraction of the $2^{15}$ Factorial 266
6.15. Constructing Other Two-Level Fractions 269
6.16. Elimination of Block Effects 271
References and Further Reading 273

Chapter 7 Additional Fractionalords and Analysis 281
7.1. Plackett and Burman Designs 281
7.2. Choosing Follow-Up Runs 294
7.3. Justifications for the Use of Fractional 303
Appendix 7A. Technical Details 305
Appendix 7B. An Approximate Partial Analysis for PB Designs 308
Appendix 7C. Hall’s Orthogonal Designs 310
References and Further Reading 313

Chapter 8 Factorial Designs and Data Transformation 317
8.1. A Two-Way (Factorial) Design 317
8.2. Simplification and Increased Sensitivity from Transformation 320
Appendix 8A. Rationale for Data Transformation 329
Appendix 8B. Bartlett’s $\chi^2$ for Testing Inhomogeneity of Variance 329
References and Further Reading 329

Chapter 9 Multiple Sources of Variation 335
9.1. Split-Plot Designs, Variance Components, and Error Transmission 335
9.2. Split-Plot Designs 335
9.3. Estimating Variance Components 345
9.4. Transmission of Error 353
References and Further Reading 359

Chapter 10 Least Squares and Why We Need Designed Experiments 363
10.1. Estimation With Least Squares 364
10.2. The Versatility of Least Squares 378
10.3. The Origins of Experimental Design 397
10.4. Nonlinear Models
Appendix 10A. Vector Representation of Statistical Concepts
Appendix 10B. Matrix Version of Least Squares
Appendix 10C. Analysis of Factorials, Botched and Otherwise
Appendix 10D. Unweighted and Weighted Least Squares
References and Further Reading

Chapter 11 Modeling, Geometry, and Experimental Design
11.1. Some Empirical Models
11.2. Some Experimental Designs and the Design Information Function
11.3. Is the Surface Sufficiently Well Estimated?
11.4. Sequential Design Strategy
11.5. Canonical Analysis
11.6. Box–Behnken Designs
References and Further Reading

Chapter 12 Some Applications of Response Surface Methods
12.1. Iterative Experimentation To Improve a Product Design
12.2. Simplification of a Response Function by Data Transformation
12.3. Detecting and Exploiting Active and Inactive Factor Spaces for Multiple-Response Data
12.4. Exploring Canonical Factor Spaces
12.5. From Empiricism to Mechanism
12.6. Uses of RSM
Appendix 12A. Average Variance of $\hat{y}$
Appendix 12B.
References and Further Reading

Chapter 13 Designing Robust Products and Processes: An Introduction
13.1. Environmental Robustness
13.2. Robustness To Component Variation
Appendix 13A. A Mathematical Formulation for Environmental Robustness
Appendix 13B. Choice of Criteria
References and Further Reading
Chapter 14 Process Control, Forecasting, and Time Series: An Introduction

14.1. Process Monitoring 565
14.2. The Exponentially Weighted Moving Average 569
14.3. The CuSum Chart 574
14.4. Process Adjustment 576
14.5. A Brief Look At Some Time Series Models and Applications 585
14.6. Using a Model to Make a Forecast 588
14.7. Intervention Analysis: A Los Angeles Air Pollution Example 593
References and Further Reading 595

Chapter 15 Evolutionary Process Operation 599

15.1. More than One Factor 602
15.2. Multiple Responses 606
15.3. The Evolutionary Process Operation Committee 607
References and Further Reading 608

Appendix Tables 611

Author Index 625

Subject Index 629