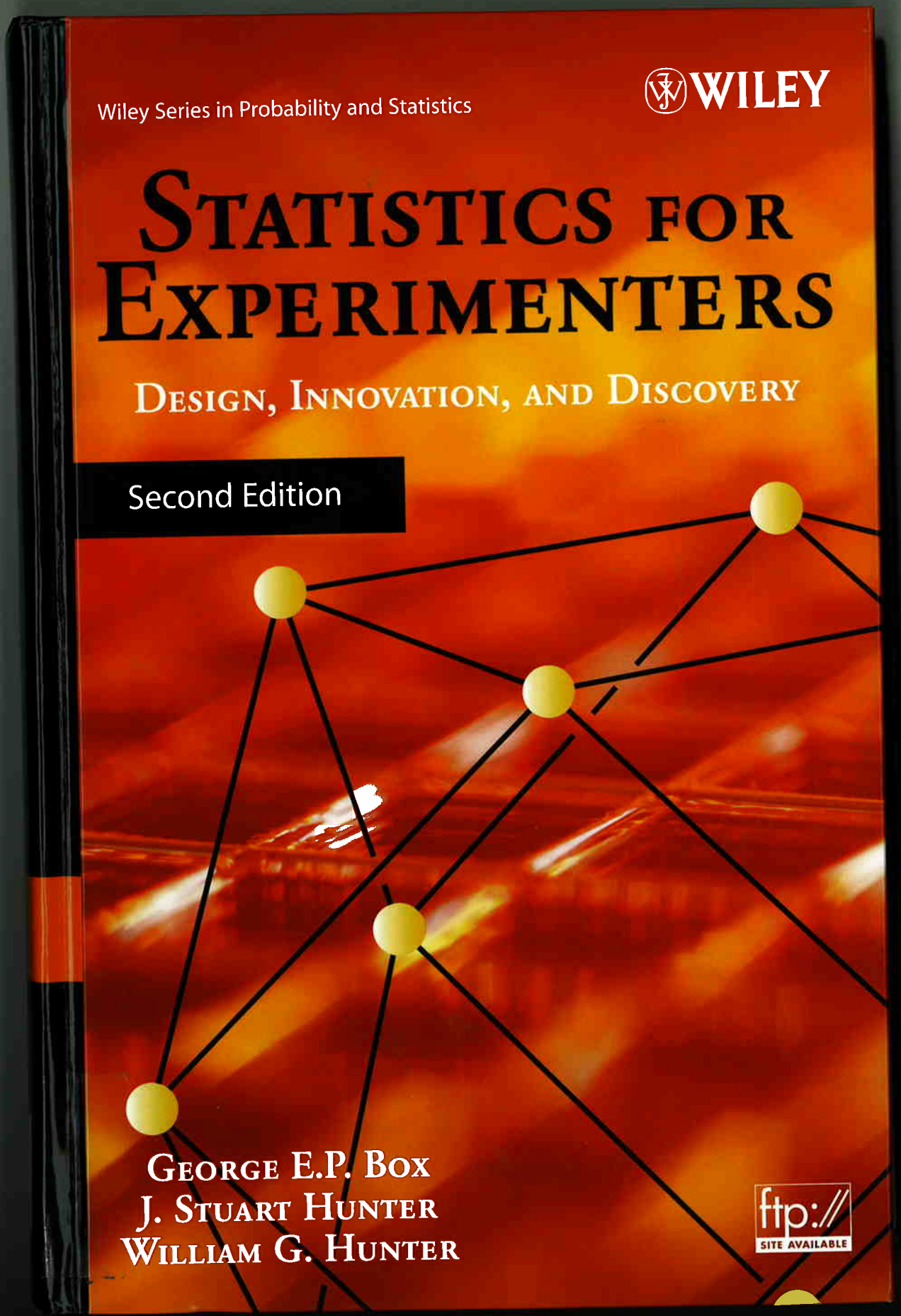
Wiley Series in Probability and Statistics
WILEY
STATISTICS FOR EXPERIMENTERS
DESIGN, INNOVATION, AND DISCOVERY
Second Edition
GEORGE E.P. BOX
J. STUART HUNTER
WILLIAM G. HUNTER
ftp://
SITE AVAILABLE

Greek Alphabet

A α	alpha	N ν	nu
B β	beta	Ξ ξ	xi
Γ γ	gamma	O o	omicron
Δ δ	delta	Π π	pi
E ε	epsilon	P ρ	rho
Z ζ	zeta	Σ σ	sigma
H η	eta	T τ	tau
Θ θ	theta	Υ υ	upsilon
I ι	iota	Φ ϕ	phi
K κ	kappa	X χ	chi
Λ λ	lambda	Ψ ψ	psi
M μ	mu	Ω ω	omega

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

Appendix 2A. Mean and Variance of Linear Combinations of Observations 57
References and Further Reading 60

Chapter 3 Comparing Two Entities: Reference Distributions, Tests, and Confidence Intervals 67

3.1. Relevant Reference Sets and Distributions 67
3.2. Randomized Paired Comparison Design: Boys' Shoes Example 81
3.3. Blocking and Randomization 92
3.4. Reprise: Comparison, Replication, Randomization, and Blocking in Simple Experiments 94
3.5. More on Significance Tests 94
3.6. Inferences About Data that are Discrete: Binomial Distribution 105
3.7. Inferences about Frequencies (Counts Per Unit): The Poisson Distribution 110
3.8. Contingency Tables and Tests of Association 112
Appendix 3A. Comparison of the Robustness of Tests to Compare Two Entities 117
Appendix 3B. Calculation of reference distribution from past data 120
References and Further Reading 123

Chapter 4 Comparing a Number of Entities, Randomized Blocks, and Latin Squares 133

4.1. Comparing k Treatments in a Fully Randomized Design 133
4.2. Randomized Block Designs 145
4.3. A Preliminary Note on Split-Plot Experiments and their Relationship to Randomized Blocks 156
4.4. More than one blocking component: Latin Squares 157
4.5. Balanced Incomplete Block Designs 162
Appendix 4A. The Rationale for the Graphical ANOVA 166
Appendix 4B. Some Useful Latin Square, Graeco–Latin Square, and Hyper-Graeco–Latin Square Designs 167
References and Further Reading 168

Chapter 5 Factorial Designs at Two Levels 173

5.1. Introduction 173

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
5.13. A 2^4 Factorial Design: Process Development Study 199
5.14. Analysis Using Normal and Lenth Plots 203
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}_{III} 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 Nodal Half Replicate of the 2^5 Factorial: Reactor Example 259
6.13. The 2_{IV}^{8-4} Nodal Sixteenth Fraction of a 2^8 Factorial 263
6.14. The 2_{III}^{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 Fractionals and Analysis 281

7.1. Plackett and Burman Designs 281
7.2. Choosing Follow-Up Runs 294
7.3. Justifications for the Use of Fractionals 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 χ_v^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 407
Appendix 10A. Vector Representation of Statistical Concepts 410
Appendix 10B. Matrix Version of Least Squares 416
Appendix 10C. Analysis of Factorials, Botched and Otherwise 418
Appendix 10D. Unweighted and Weighted Least Squares 420
References and Further Reading 424

Chapter 11 Modeling, Geometry, and Experimental Design 437

11.1. Some Empirical Models 441
11.2. Some Experimental Designs and the Design Information Function 447
11.3. Is the Surface Sufficiently Well Estimated? 453
11.4. Sequential Design Strategy 454
11.5. Canonical Analysis 461
11.6. Box–Behnken Designs 475
References and Further Reading 483

Chapter 12 Some Applications of Response Surface Methods 489

12.1. Iterative Experimentation To Improve a Product Design 489
12.2. Simplification of a Response Function by Data Transformation 503
12.3. Detecting and Exploiting Active and Inactive Factor Spaces for Multiple-Response Data 509
12.4. Exploring Canonical Factor Spaces 513
12.5. From Empiricism to Mechanism 518
12.6. Uses of RSM 526
Appendix 12A. Average Variance of $\hat{y}$ 526
Appendix 12B. 528
References and Further Reading 530

Chapter 13 Designing Robust Products and Processes: An Introduction 539

13.1. Environmental Robustness 539
13.2. Robustness To Component Variation 549
Appendix 13A. A Mathematical Formulation for Environmental Robustness 556
Appendix 13B. Choice of Criteria 558
References and Further Reading 559

Chapter 14 Process Control, Forecasting, and Time Series: An Introduction 565

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 **621**

Subject Index **623**

Preface to the Second Edition

In rewriting this book, we have deeply felt the loss of our dear friend and colleague Bill Hunter (William G. Hunter, 1937–1986). We miss his ideas, his counsel, and his cheerful encouragement. We believe, however, that his spirit has been with us and that he would be pleased with what we have done.

The objectives for this revised edition of *Statistics for Experimenters* remain the same as those for the first:

1. To make available to experimenters scientific and statistical tools that can greatly catalyze innovation, problem solving, and discovery.
2. To illustrate how these tools may be used by and with subject matter specialists as their investigations proceed.

Developments that would have delighted Bill are the receptive atmosphere these techniques now encounter in industry and the present universal availability of very fast computers, allowing* where necessary, the ready use of computationally intensive methods.

Under such banners as "Six Sigma," management has realized the importance of training its work forces in the arts of economic investigation. With this democratization of the scientific method, many more people are being found with creative ability and unrealized aptitude for problem solving and discovery. Also, the "team idea" not only accelerates improvement but identifies such natural leaders of innovation and can allow them to lead. To make such initiatives possible the modern philosophy and methods of process improvement must be taught at all levels of an organization. We believe both trainers and trainees engaged in such efforts will find this book helpful. Also based on a long experience, we

*All the computations in this book can be done with the statistical language R (R Development Core Team, 2004), available at CRAN (http://cran.R-project.org). Functions for displaying anova and lambda plots, for Bayesian screening and model building are included in the BHH2 and BsMD R-packages and available at CRAN under contributed packages. There is as well commercial software, such as the SCA Statistical System, which some readers will find easier to use.

The classic adapted to modern times

For many years, the *First Edition* of *Statistics for Experimenters* has been a premier guide and reference for the application of statistical methods, especially as applied to experimental design. Rewritten and updated, this new edition of *Statistics for Experimenters* adopts the same approach as th[illegible]rating thoroughly worked examples, readily understood graphics, and the appropriate [illegible]tion, problem solving, and discovery, the *Second Edition* provides experimenters with t[illegible]eded to maximize the knowledge gained from research data. The authors' practical app[illegible]eds to be solved and then illustrates the statistical methods best utilized in all stages of design and analysis.

HKU Libraries
X46394234

Providing even greater accessibility for its users, the *Second Edition* reflects the changes in techniques and technologies since the publication of the classic *First Edition*.

Among the new topics included are:

- Graphical analysis of variance
- Computer analysis of complex designs
- Simplification by transformation
- Hands-on experimentation using response surface methods
- Further development of robust product and process design using split-plot arrangements and minimization of error transmission
- Introduction to process control, forecasting, and time series
- Illustrations demonstrating how multiresponse problems can be solved using the concepts of active and inert factor spaces and canonical spaces
- Bayesian approaches to model selection and sequential experimentation
- Applications for Six Sigma initiatives in a variety of disciplines
- An appendix featuring quaquaversal quotes from a variety of sources ranging from noted statisticians and scientists to famous philosophers that embellish key concepts and enliven the learning process

Computations in the *Second Edition* can be done utilizing the statistical language R. Functions for displaying ANOVA and lambda plots, Bayesian screening, and model building are all included, and R packages are available on a related FTP site. These topics can also be applied utilizing easy-to-use commercial software packages.

Complete with applications covering the physical, engineering, biological, and social sciences, *Statistics for Experimenters* is designed for all individuals who must use statistical approaches to conduct an experiment. Experimenters need only a basic understanding of mathematics to master all the statistical methods presented. This text is an essential reference for all researchers and an invaluable course book for undergraduate and graduate students.

GEORGE E.P. BOX, PhD, is Ronald Aylmer Fisher Professor Emeritus of Statistics at the University of Wisconsin–Madison. He is a Fellow of the American Academy of Arts and Sciences and a recipient of the Samuel S. Wilks Memorial Medal of the American Statistical Association, the Shewhart Medal of the American Society for Quality, and the Guy Medal in Silver of the Royal Statistical Society.

J. STUART HUNTER, PhD, is Professor Emeritus of Civil Engineering at Princeton University. Dr. Hunter has served as consultant to many industries and government agencies. He has been a staff member of the National Academy of Sciences, Committee on National Statistics, and statistician in residence at the University of Wisconsin, and is the founding editor of *Technometrics*.

The late **WILLIAM G. HUNTER, PhD,** was Professor Emeritus of Statistics and Engineering at the University of Wisconsin–Madison.

Cover Design: Abbie Enneking
Cover Photograph: Ryan McVay/Getty Images

Subscribe to our free Statistics eNewsletter at
www.wiley.com/enewsletters

Visit www.wiley.com/statistics

WILEY-INTERSCIENCE
wiley.com

ISBN 0-471-71813-0
9 780471 718130 90000