

Contents

1	Principles and Techniques	1
1.1	Design: Basic Principles and Techniques	1
1.1.1	The Art of Experimentation	1
1.1.2	Replication	2
1.1.3	Blocking	2
1.1.4	Randomization	3
1.2	Analysis: Basic Principles and Techniques	4
2	Planning Experiments	7
2.1	Introduction	7
2.2	A Checklist for Planning Experiments	7
2.3	A Real Experiment—Cotton-Spinning Experiment	13
2.4	Some Standard Experimental Designs	16
2.4.1	Completely Randomized Designs	17
2.4.2	Block Designs	17
2.4.3	Designs with Two or More Blocking Factors	17
2.4.4	Split-Plot Designs	19
2.5	More Real Experiments	20
2.5.1	Soap Experiment	20
2.5.2	Battery Experiment	24
2.5.3	Cake-Baking Experiment	27
	Exercises	29
3	Designs with One Source of Variation	31
3.1	Introduction	31
3.2	Randomization	31
3.3	Model for a Completely Randomized Design	32
3.4	Estimation of Parameters	34
3.4.1	Estimable Functions of Parameters	34
3.4.2	Notation	34
3.4.3	Obtaining Least Squares Estimates	35
3.4.4	Properties of Least Squares Estimators	37
3.4.5	Estimation of σ^2	39
3.4.6	Confidence Bound for σ^2	39
3.5	One-Way Analysis of Variance	41
3.5.1	Testing Equality of Treatment Effects	41
3.5.2	Use of p -Values	45

3.6	Sample Sizes	45
3.6.1	Expected Mean Squares for Treatments	46
3.6.2	Sample Sizes Using Power of a Test	47
3.7	A Real Experiment—Soap Experiment, Continued	49
3.7.1	Checklist, Continued	50
3.7.2	Data Collection and Analysis	50
3.7.3	Discussion by the Experimenter.	52
3.7.4	Further Observations by the Experimenter	52
3.8	Using SAS Software	52
3.8.1	Randomization	52
3.8.2	Analysis of Variance	54
3.8.3	Calculating Sample Size Using Power of a Test.	56
3.9	Using R Software	57
3.9.1	Randomization	59
3.9.2	Reading and Plotting Data	60
3.9.3	Analysis of Variance	62
3.9.4	Calculating Sample Size Using Power of a Test.	64
	Exercises.	65
4	Inferences for Contrasts and Treatment Means	69
4.1	Introduction	69
4.2	Contrasts	69
4.2.1	Pairwise Comparisons	70
4.2.2	Treatment Versus Control.	71
4.2.3	Difference of Averages.	72
4.2.4	Trends	72
4.3	Individual Contrasts and Treatment Means	74
4.3.1	Confidence Interval for a Single Contrast	74
4.3.2	Confidence Interval for a Single Treatment Mean	76
4.3.3	Hypothesis Test for a Single Contrast or Treatment Mean	77
4.3.4	Equivalence of Tests and Confidence Intervals (Optional)	79
4.4	Methods of Multiple Comparisons.	81
4.4.1	Multiple Confidence Intervals	81
4.4.2	Bonferroni Method for Preplanned Comparisons.	83
4.4.3	Scheffé Method of Multiple Comparisons.	85
4.4.4	Tukey Method for All Pairwise Comparisons.	87
4.4.5	Dunnnett Method for Treatment-Versus- Control Comparisons	90
4.4.6	Combination of Methods	92
4.4.7	Methods Not Controlling Experimentwise Error Rate	92
4.5	Sample Sizes	92

4.6	Using SAS Software	94
4.6.1	Inferences on Individual Contrasts	94
4.6.2	Multiple Comparisons	95
4.7	Using R Software	96
4.7.1	Inferences on Individual Contrasts	97
4.7.2	Multiple Comparisons	99
	Exercises.	100
5	Checking Model Assumptions	103
5.1	Introduction	103
5.2	Strategy for Checking Model Assumptions.	103
5.2.1	Residuals	104
5.2.2	Residual Plots	104
5.3	Checking the Fit of the Model	106
5.4	Checking for Outliers	107
5.5	Checking Independence of the Error Terms	108
5.6	Checking the Equal Variance Assumption	110
5.6.1	Detection of Unequal Variances	110
5.6.2	Data Transformations to Equalize Variances	112
5.6.3	Analysis with Unequal Error Variances	115
5.7	Checking the Normality Assumption	117
5.8	Using SAS Software	119
5.8.1	Residual Plots	119
5.8.2	Transforming the Data	123
5.8.3	Implementing Satterthwaite's Method.	124
5.9	Using R Software	125
5.9.1	Residual Plots	125
5.9.2	Transforming the Data	129
5.9.3	Implementing Satterthwaite's Method.	130
	Exercises.	132
6	Experiments with Two Crossed Treatment Factors	139
6.1	Introduction	139
6.2	Models and Factorial Effects	139
6.2.1	The Meaning of Interaction.	139
6.2.2	Models for Two Treatment Factors	142
6.2.3	Checking the Assumptions on the Model	143
6.3	Contrasts	144
6.3.1	Contrasts for Main Effects and Interactions.	144
6.3.2	Writing Contrasts as Coefficient Lists	146
6.4	Analysis of the Two-Way Complete Model	149
6.4.1	Least Squares Estimators for the Two-Way Complete Model	149
6.4.2	Estimation of σ^2 for the Two-Way Complete Model	151
6.4.3	Multiple Comparisons for the Complete Model	152
6.4.4	Analysis of Variance for the Complete Model	155

6.5	Analysis of the Two-Way Main-Effects Model	161
6.5.1	Least Squares Estimators for the Main-Effects Model	161
6.5.2	Estimation of σ^2 in the Main-Effects Model	165
6.5.3	Multiple Comparisons for the Main-Effects Model	166
6.5.4	Unequal Variances	168
6.5.5	Analysis of Variance for Equal Sample Sizes	168
6.5.6	Model Building	170
6.6	Calculating Sample Sizes	171
6.7	Small Experiments	171
6.7.1	One Observation Per Cell	171
6.7.2	Analysis Based on Orthogonal Contrasts	172
6.7.3	Tukey's Test for Additivity	175
6.7.4	A Real Experiment—Air Velocity Experiment	176
6.8	Using SAS Software	177
6.8.1	Analysis of Variance	177
6.8.2	Contrasts and Multiple Comparisons	180
6.8.3	Plots	182
6.8.4	One Observation Per Cell	183
6.9	Using R Software	184
6.9.1	Analysis of Variance	186
6.9.2	Contrasts and Multiple Comparisons	187
6.9.3	Plots	191
6.9.4	One Observation Per Cell	192
	Exercises	193
7	Several Crossed Treatment Factors	201
7.1	Introduction	201
7.2	Models and Factorial Effects	201
7.2.1	Models	201
7.2.2	The Meaning of Interaction	202
7.2.3	Separability of Factorial Effects	205
7.2.4	Estimation of Factorial Contrasts	206
7.3	Analysis—Equal Sample Sizes	209
7.4	A Real Experiment—Popcorn—Microwave Experiment	213
7.5	One Observation per Cell	219
7.5.1	Analysis Assuming that Certain Interaction Effects are Negligible	219
7.5.2	Analysis Using Half-Normal Probability Plot of Effect Estimates	221
7.5.3	Analysis Using Confidence Intervals	223

7.6	Using SAS Software	225
7.6.1	Half-Normal Probability Plots of Contrast Estimates	225
7.6.2	Voss–Wang Confidence Interval Method	227
7.6.3	Experiments with Empty Cells	228
7.7	Using R Software	231
7.7.1	Half-Normal Probability Plots of Contrast Estimates	231
7.7.2	Voss–Wang Confidence Interval Method	232
7.7.3	Experiments with Empty Cells	234
	Exercises.	237
8	Polynomial Regression	249
8.1	Introduction	249
8.2	Models	250
8.3	Least Squares Estimation (Optional)	253
8.3.1	Normal Equations	253
8.3.2	Least Squares Estimates for Simple Linear Regression	254
8.4	Test for Lack of Fit.	254
8.5	Analysis of the Simple Linear Regression Model	257
8.6	Analysis of Polynomial Regression Models	260
8.6.1	Analysis of Variance	260
8.6.2	Confidence Intervals	262
8.7	Orthogonal Polynomials and Trend Contrasts (Optional)	263
8.7.1	Simple Linear Regression	263
8.7.2	Quadratic Regression	265
8.7.3	Comments	266
8.8	A Real Experiment—Bean-Soaking Experiment	266
8.8.1	Checklist	267
8.8.2	One-Way Analysis of Variance and Multiple Comparisons	270
8.8.3	Regression Analysis.	271
8.9	Using SAS Software	273
8.10	Using R Software	276
	Exercises.	281
9	Analysis of Covariance	285
9.1	Introduction	285
9.2	Models	285
9.2.1	Checking Model Assumptions and Equality of Slopes	287
9.2.2	Model Extensions	287
9.3	Least Squares Estimates	287
9.3.1	Normal Equations (Optional).	287
9.3.2	Least Squares Estimates and Adjusted Treatment Means	288
9.4	Analysis of Covariance	289

9.5	Treatment Contrasts and Confidence Intervals.	293
9.5.1	Individual Confidence Intervals	293
9.5.2	Multiple Comparisons	294
9.6	Using SAS Software	296
9.7	Using R Software	299
	Exercises.	301
10	Complete Block Designs	305
10.1	Introduction	305
10.2	Blocks, Noise Factors or Covariates?	305
10.3	Design Issues	306
10.3.1	Block Sizes.	306
10.3.2	Complete Block Design Definitions	307
10.3.3	The Randomized Complete Block Design.	308
10.3.4	The General Complete Block Design	309
10.3.5	How Many Observations?	309
10.4	Analysis of Randomized Complete Block Designs	310
10.4.1	Model and Analysis of Variance	310
10.4.2	Multiple Comparisons	313
10.5	A Real Experiment—Cotton-Spinning Experiment	314
10.5.1	Design Details.	314
10.5.2	Sample-Size Calculation	315
10.5.3	Analysis of the Cotton-Spinning Experiment.	315
10.6	Analysis of General Complete Block Designs.	318
10.6.1	Model and Analysis of Variance	318
10.6.2	Multiple Comparisons for the General Complete Block Design	320
10.6.3	Sample-Size Calculations	322
10.7	Checking Model Assumptions	323
10.8	Factorial Experiments	324
10.9	Using SAS Software	327
10.10	Using R Software	331
	Exercises.	336
11	Incomplete Block Designs	349
11.1	Introduction	349
11.2	Design Issues	349
11.2.1	Block Sizes.	349
11.2.2	Design Plans and Randomization.	350
11.2.3	Estimation of Contrasts	351
11.3	Some Special Incomplete Block Designs	352
11.3.1	Balanced Incomplete Block Designs	352
11.3.2	Group Divisible Designs.	354
11.3.3	Cyclic Designs	355
11.4	Analysis of General Incomplete Block Designs.	356
11.4.1	Contrast Estimators and Multiple Comparisons.	356
11.4.2	Analysis of Variance	358

11.4.3	Analysis of Balanced Incomplete Block Designs	359
11.4.4	A Real Experiment—Detergent Experiment	360
11.4.5	Analysis of Group Divisible Designs	365
11.4.6	Analysis of Cyclic Designs	366
11.5	A Real Experiment—Plasma Experiment	366
11.6	Sample Sizes	371
11.7	Factorial Experiments	372
11.7.1	Factorial Structure	372
11.8	Using SAS Software	376
11.8.1	Generation of Efficient Block Designs	376
11.8.2	Analysis of Variance, Contrasts, and Multiple Comparisons	378
11.8.3	Plots	381
11.9	Using R Software	382
11.9.1	Generating Efficient Incomplete Block Designs	382
11.9.2	Analysis of Variance, Contrasts, and Multiple Comparisons	384
11.9.3	Plots	387
	Exercises	389
12	Designs with Two Blocking Factors	399
12.1	Introduction	399
12.2	Design Issues	400
12.2.1	Selection and Randomization of Row–Column Designs	400
12.2.2	Latin Square Designs	401
12.2.3	Youden Designs	403
12.2.4	Cyclic and Other Row–Column Designs	404
12.3	Analysis of Row–Column Designs	405
12.3.1	Model for a Row–Column Design	405
12.3.2	Analysis of Variance for a Row–Column Design	406
12.3.3	Confidence Intervals and Multiple Comparisons	407
12.4	Analysis of Latin Square Designs	408
12.4.1	Analysis of Variance for Latin Square Designs	408
12.4.2	Confidence Intervals for Latin Square Designs	409
12.4.3	How Many Observations?	411
12.5	Analysis of Youden Designs	412
12.5.1	Analysis of Variance for Youden Designs	412
12.5.2	Confidence Intervals for Youden Designs	413
12.5.3	How Many Observations?	413
12.6	Checking the Assumptions on the Model	414
12.7	Factorial Experiments in Row–Column Designs	416

12.8	Using SAS Software	416
12.8.1	Factorial Model	417
12.8.2	Plots	419
12.9	Using R Software	420
12.9.1	Factorial Model	423
12.9.2	Plots	424
	Exercises	425
13	Confounded Two-Level Factorial Experiments	433
13.1	Introduction	433
13.2	Single Replicate Factorial Experiments	433
13.2.1	Coding and Notation	433
13.2.2	Confounding	434
13.2.3	Analysis	434
13.3	Confounding Using Contrasts	435
13.3.1	Contrasts	435
13.3.2	Experiments in Two Blocks	436
13.3.3	Experiments in Four Blocks	441
13.3.4	Experiments in Eight Blocks	443
13.3.5	Experiments in More Than Eight Blocks	443
13.4	Confounding Using Equations	444
13.4.1	Experiments in Two Blocks	444
13.4.2	Experiments in More Than Two Blocks	445
13.5	A Real Experiment—Mangold Experiment	447
13.6	Plans for Confounded 2^p Experiments	451
13.7	Multireplicate Designs	451
13.8	Complete Confounding: Repeated Single-Replicate Designs	452
13.8.1	A Real Experiment—Decontamination Experiment	452
13.9	Partial Confounding	455
13.10	Comparing the Multireplicate Designs	458
13.11	Using SAS Software	461
13.12	Using R Software	463
	Exercises	465
14	Confounding in General Factorial Experiments	473
14.1	Introduction	473
14.2	Confounding with Factors at Three Levels	473
14.2.1	Contrasts	473
14.2.2	Confounding Using Contrasts	474
14.2.3	Confounding Using Equations	475
14.2.4	A Real Experiment—Dye Experiment	478
14.2.5	Plans for Confounded 3^p Experiments	481
14.3	Designing Using Pseudofactors	482
14.3.1	Confounding in 4^p Experiments	482
14.3.2	Confounding in $2^p \times 4^q$ Experiments	483
14.4	Designing Confounded Asymmetric Experiments	483

14.5	Using SAS Software	486
14.6	Using R Software	489
	Exercises.	491
15	Fractional Factorial Experiments	495
15.1	Introduction	495
15.2	Fractions from Block Designs; Factors with 2 Levels	495
15.2.1	Half-Fractions of 2^p Experiments; 2^{p-1} Experiments.	495
15.2.2	Resolution and Notation	498
15.2.3	A Real Experiment—Soup Experiment.	499
15.2.4	Quarter-Fractions of 2^p Experiments; 2^{p-2} Experiments.	501
15.2.5	Smaller Fractions of 2^p Experiments	505
15.3	Fractions from Block Designs; Factors with 3 Levels	507
15.3.1	One-Third Fractions of 3^p Experiments; 3^{p-1} Experiments.	507
15.3.2	One-Ninth Fractions of 3^p Experiments; 3^{p-2} Experiments.	511
15.4	Fractions from Block Designs; Other Experiments.	511
15.4.1	$2^p \times 4^q$ Experiments	511
15.4.2	$2^p \times 3^q$ Experiments	512
15.5	Blocked Fractional Factorial Experiments.	513
15.6	Fractions from Orthogonal Arrays.	516
15.6.1	2^p Orthogonal Arrays.	516
15.6.2	$2^p \times 4^q$ Orthogonal Arrays	521
15.6.3	3^p Orthogonal Arrays.	522
15.7	Design for the Control of Noise Variability	523
15.7.1	A Real Experiment—Inclinometer Experiment	525
15.8	Small Screening Designs: Orthogonal Main Effect Plans	529
15.8.1	Saturated Designs	529
15.8.2	Supersaturated Designs.	533
15.8.3	Saturated Orthogonal Main Effect Plans Plus Interactions	536
15.8.4	Definitive Screening Designs	537
15.9	Using SAS Software	538
15.9.1	Fractional Factorials.	538
15.9.2	Design for the Control of Noise Variability.	539
15.9.3	Analysis of Small Screening Designs	542
15.10	Using R Software	543
15.10.1	Fractional Factorials.	543
15.10.2	Design for the Control of Noise Variability.	546
15.10.3	Analysis of Small Screening Designs	547
	Exercises.	549

16	Response Surface Methodology	565
16.1	Introduction	565
16.2	First-Order Designs and Analysis	567
16.2.1	Models	567
16.2.2	Standard First-Order Designs	568
16.2.3	Least Squares Estimation	569
16.2.4	Checking Model Assumptions	570
16.2.5	Analysis of Variance	570
16.2.6	Tests for Lack of Fit	572
16.2.7	Path of Steepest Ascent	576
16.3	Second-Order Designs and Analysis	577
16.3.1	Models and Designs	577
16.3.2	Central Composite Designs	578
16.3.3	Generic Test for Lack of Fit of the Second-Order Model	581
16.3.4	Analysis of Variance for a Second-Order Model	581
16.3.5	Canonical Analysis of a Second-Order Model	583
16.4	Properties of Second-Order Designs: CCDs	585
16.4.1	Rotatability	585
16.4.2	Orthogonality	586
16.4.3	Orthogonal Blocking	587
16.5	A Real Experiment: Flour Production Experiment, Continued	589
16.6	Box–Behnken Designs	592
16.7	Using SAS Software	594
16.7.1	Analysis of a Standard First-Order Design	594
16.7.2	Analysis of a Second-Order Design	597
16.8	Using R Software	599
16.8.1	Analysis of a Standard First-Order Design	599
16.8.2	Analysis of a Second-Order Design	602
16.8.3	Generating Designs	606
	Exercises	608
17	Random Effects and Variance Components	615
17.1	Introduction	615
17.2	Some Examples	615
17.3	One Random Effect	618
17.3.1	The Random-Effects One-Way Model	618
17.3.2	Estimation of σ^2	619
17.3.3	Estimation of σ_T^2	620
17.3.4	Testing Equality of Treatment Effects	622
17.3.5	Confidence Intervals for Variance Components	625
17.4	Sample Sizes for an Experiment with One Random Effect	628
17.5	Checking Assumptions on the Model	631

17.6	Two or More Random Effects	632
17.6.1	Models and Examples	632
17.6.2	Checking Model Assumptions	634
17.6.3	Estimation of σ^2	634
17.6.4	Estimation of Variance Components	635
17.6.5	Confidence Intervals for Variance Components	637
17.6.6	Hypothesis Tests for Variance Components	640
17.6.7	Sample Sizes	642
17.7	Mixed Models	642
17.7.1	Expected Mean Squares and Hypothesis Tests	643
17.7.2	Confidence Intervals in Mixed Models	645
17.8	Rules for Analysis of Random-Effects and Mixed Models	647
17.8.1	Rules—Equal Sample Sizes	647
17.8.2	Controversy (Optional)	648
17.9	Block Designs and Random Block Effects	649
17.10	Using SAS Software	652
17.10.1	Checking Assumptions on the Model	652
17.10.2	Estimation and Hypothesis Testing	654
17.10.3	Sample Size Calculations	657
17.11	Using R Software	659
17.11.1	Checking Assumptions on the Model	659
17.11.2	Estimation and Hypothesis Testing	661
17.11.3	Sample Size Calculations	664
	Exercises	665
18	Nested Models	671
18.1	Introduction	671
18.2	Examples and Models	671
18.3	Analysis of Nested Fixed Effects	674
18.3.1	Least Squares Estimates	674
18.3.2	Estimation of σ^2	675
18.3.3	Confidence Intervals	675
18.3.4	Hypothesis Testing	676
18.4	Analysis of Nested Random Effects	679
18.4.1	Expected Mean Squares	679
18.4.2	Estimation of Variance Components	682
18.4.3	Hypothesis Testing	683
18.4.4	Some Examples	683
18.5	Using SAS Software	687
18.5.1	Voltage Experiment	687

18.6	Using R Software	691
18.6.1	Voltage Experiment	692
18.6.2	Analysis Using Least Squares Estimates and aov	694
18.6.3	Analysis Using Restricted Maximum Likelihood Estimation	695
	Exercises	695
19	Split-Plot Designs	703
19.1	Introduction	703
19.2	Designs and Models	703
19.3	Analysis of a Split-Plot Design with Complete Blocks	705
19.3.1	Split-Plot Analysis	706
19.3.2	Whole-Plot Analysis	707
19.3.3	Contrasts Within and Between Whole Plots	708
19.3.4	A Real Experiment—Oats Experiment	709
19.4	Split-Split-Plot Designs	711
19.5	Split-Plot Confounding	713
19.6	A Real Experiment—UAV Experiment	713
19.6.1	Analysis of Variance	715
19.6.2	Multiple Comparisons	716
19.7	A Real Experiment—Mobile Computing Field Study	717
19.7.1	Analysis of Variance	719
19.7.2	Multiple Comparisons	721
19.7.3	Analysis as a Split-Split-Plot Design	721
19.7.4	Design Construction	724
19.8	Using SAS Software	727
19.8.1	The Analysis of Variance Approach	727
19.8.2	Simple Contrasts	732
19.8.3	The Restricted Maximum Likelihood Approach	733
19.8.4	Recovery of Inter-block Information	736
19.8.5	ReML and Unbalanced Data	742
19.9	Using R Software	745
19.9.1	The Analysis of Variance Approach	745
19.9.2	Simple Contrasts	746
19.9.3	The Restricted Maximum Likelihood Approach	747
19.9.4	Recovery of Inter-block Information	753
19.9.5	ReML and Unbalanced Data	756
	Exercises	758
20	Computer Experiments	765
20.1	Introduction	765
20.2	Models for Computer Experiments	767
20.3	Gaussian Stochastic Process Model	768

20.4	Design.	772
20.4.1	Space-Filling and Non-collapsing Designs . . .	772
20.4.2	Construction of Latin Hypercube Designs. . . .	774
20.5	A Real Experiment—Neuron Experiment.	776
20.6	Using SAS Software	778
20.6.1	Maximin Latin Hypercube Designs	779
20.6.2	Fitting the GaSP Model	781
20.7	Using R Software	784
20.7.1	Maximin Latin Hypercube Designs	784
20.7.2	Fitting the GaSP Model	785
	Exercises.	786
	Appendix A: Tables	793
	Bibliography	821
	Index of Authors	825
	Index of Experiments	829
	Index of Subjects	833

One of the first questions facing an experimenter is, "How many observations do I need to take?" or alternatively, "Given my limited budget / cost, can I gain as much information as possible?" There are not questions that can be answered in a couple of sentences. They are, however, questions that are central to the material in this book. As a first step towards choosing an answer, the experimenter will ask further questions, such as, "What is the main purpose of running this experiment?" and "What do I hope to be able to learn?"

Typically, an experiment may be run for one or more of the following reasons:

- (a) to describe the principal causes of variation in a measured response;
- (b) to find the conditions that give rise to a maximum or minimum response;
- (c) to compare the responses achieved at different settings of controllable variables;
- (d) to obtain a mathematical model in order to predict future responses.

Observations can be collected from observational studies as well as from experiments, but only an experiment allows conclusions to be drawn about causal effect. For example, consider the following situation:

The output from each machine on a factory floor is constantly monitored by an sophisticated manufacturing company. Suppose that in a particular factory, the output from a particular machine is consistently of low quality. What should the managers do? They could conclude that the machine needs replacing and pay out a large sum of money for a new one. They could decide that the machine operator is at fault and dismiss him or her. They could conclude that the humidity in that part of the factory is too high and install a new air conditioning system. In other words, the machine output has been observed under the current operating conditions (an observational study), and although it has been very effective in showing the management that a problem exists, it has given them very little idea about the cause of the poor quality.

It would actually be a simple matter to determine or rule out some of the potential causes. For example, the question about the operator could be answered by moving all the operations from machine to machine over several days. If the poor output follows the operator, then it is safe to conclude that the operator is the cause. If the poor output remains with the original machine, then the operator is