Contents

Preface *page* xvii

PART I INTRODUCTION

1 Causality: The Basic Framework 3
- 1.1 Introduction 3
- 1.2 Potential Outcomes 3
- 1.3 Definition of Causal Effects 5
- 1.4 Causal Effects in Common Usage 7
- 1.5 Learning about Causal Effects: Multiple Units 8
- 1.6 The Stable Unit Treatment Value Assumption 9
- 1.7 The Assignment Mechanism: An Introduction 13
- 1.8 Attributes, Pre-Treatment Variables, or Covariates 15
- 1.9 Potential Outcomes and Lord's Paradox 16
- 1.10 Causal Estimands 18
- 1.11 Structure of the Book 20
- 1.12 Samples, Populations, and Super-Populations 20
- 1.13 Conclusion 21
- Notes 21

2 A Brief History of the Potential Outcomes Approach to Causal Inference 23
- 2.1 Introduction 23
- 2.2 Potential Outcomes and the Assignment Mechanism before Neyman 24
- 2.3 Neyman's (1923) Potential Outcome Notation in Randomized Experiments 25
- 2.4 Earlier Hints for Physical Randomizing 26
- 2.5 Fisher's (1925) Proposal to Randomize Treatments to Units 26
- 2.6 The Observed Outcome Notation in Observational Studies for Causal Effects 27
- 2.7 Early Uses of Potential Outcomes in Observational Studies in Social Sciences 28

2.8 Potential Outcomes and the Assignment Mechanism in Observational Studies: Rubin (1974) 29
Notes 30

3 A Classification of Assignment Mechanisms 31
3.1 Introduction 31
3.2 Notation 33
3.3 Assignment Probabilities 34
3.4 Restrictions on the Assignment Mechanism 37
3.5 Assignment Mechanisms and Super-Populations 39
3.6 Randomized Experiments 40
3.7 Observational Studies: Regular Assignment Mechanisms 41
3.8 Observational Studies: Irregular Assignment Mechanisms 42
3.9 Conclusion 43
Notes 43

PART II CLASSICAL RANDOMIZED EXPERIMENTS

4 A Taxonomy of Classical Randomized Experiments 47
4.1 Introduction 47
4.2 Notation 48
4.3 Bernoulli Trials 48
4.4 Completely Randomized Experiments 50
4.5 Stratified Randomized Experiments 51
4.6 Paired Randomized Experiments 52
4.7 Discussion 53
4.8 Conclusion 55
Notes 56

5 Fisher's Exact P-Values for Completely Randomized Experiments 57
5.1 Introduction 57
5.2 The Paul et al. Honey Experiment Data 59
5.3 A Simple Example with Six Units 59
5.4 The Choice of Null Hypothesis 63
5.5 The Choice of Statistic 64
5.6 A Small Simulation Study 72
5.7 Interval Estimates Based on Fisher P-Value Calculations 74
5.8 Computation of P-Values 75
5.9 Fisher Exact P-Values with Covariates 78
5.10 Fisher Exact P-Values for the Honey Data 80
5.11 Conclusion 81
Notes 81

6 Neyman's Repeated Sampling Approach to Completely Randomized Experiments 83
6.1 Introduction 83
6.2 The Duflo-Hanna-Ryan Teacher-Incentive Experiment Data 84
6.3 Unbiased Estimation of the Average Treatment Effect 85

	6.4	The Sampling Variance of the Neyman Estimator	87
	6.5	Estimating the Sampling Variance	92
	6.6	Confidence Intervals and Testing	95
	6.7	Inference for Population Average Treatment Effects	98
	6.8	Neyman's Approach with Covariates	101
	6.9	Results for the Duflo-Hanna-Ryan Teacher-Incentive Data	102
	6.10	Conclusion	104
		Notes	104
		Appendix A Sampling Variance Calculations	105
		Appendix B Random Sampling from a Super-Population	109
7	**Regression Methods for Completely Randomized Experiments**		113
	7.1	Introduction	113
	7.2	The LRC-CPPT Cholesterol Data	115
	7.3	The Super-Population Average Treatment Effects	116
	7.4	Linear Regression with No Covariates	118
	7.5	Linear Regression with Additional Covariates	122
	7.6	Linear Regression with Covariates and Interactions	125
	7.7	Transformations of the Outcome Variable	127
	7.8	The Limits on Increases in Precision Due to Covariates	128
	7.9	Testing for the Presence of Treatment Effects	129
	7.10	Estimates for LRC-CPPT Cholesterol Data	131
	7.11	Conclusion	133
		Notes	134
		Appendix	135
8	**Model-Based Inference for Completely Randomized Experiments**		141
	8.1	Introduction	141
	8.2	The Lalonde NSW Experimental Job-Training Data	144
	8.3	A Simple Example: Naive and More Sophisticated Approaches to Imputation	146
	8.4	Bayesian Model-Based Imputation in the Absence of Covariates	150
	8.5	Simulation Methods in the Model-Based Approach	163
	8.6	Dependence between Potential Outcomes	165
	8.7	Model-Based Imputation with Covariates	169
	8.8	Super-Population Average Treatment Effects	171
	8.9	A Frequentist Perspective	172
	8.10	Model-Based Estimates of the Effect of the NSW Program	174
	8.11	Conclusion	177
		Notes	177
		Appendix A Posterior Distributions for Normal Models	178
		Appendix B Analytic Derivations with Known Covariance Matrix	181
9	**Stratified Randomized Experiments**		187
	9.1	Introduction	187
	9.2	The Tennesee Project Star Data	188

9.3 The Structure of Stratified Randomized Experiments 189
9.4 Fisher's Exact P-Values in Stratified Randomized Experiments 192
9.5 The Analysis of Stratified Randomized Experiments from Neyman's Repeated Sampling Perspective 201
9.6 Regression Analysis of Stratified Randomized Experiments 205
9.7 Model-Based Analysis of Stratified Randomized Experiments 207
9.8 Design Issues: Stratified versus Completely Randomized Experiments 211
9.9 Conclusion 212
Notes 212
Appendix A: Student-Level Analyses 213
Appendix B: Proofs of Theorems 9.1 and 9.2 214

10 Pairwise Randomized Experiments 219
10.1 Introduction 219
10.2 The Children's Television Workshop Experiment Data 220
10.3 Pairwise Randomized Experiments 220
10.4 Fisher's Exact P-Values in Pairwise Randomized Experiments 222
10.5 The Analysis of Pairwise Randomized Experiments from Neyman's Repeated Sampling Perspective 224
10.6 Regression-Based Analysis of Pairwise Randomized Experiments 229
10.7 Model-Based Analysis of Pairwise Randomized Experiments 231
10.8 Conclusion 233
Notes 234
Appendix: Proofs 234

11 Case Study: An Experimental Evaluation of a Labor Market Program 240
11.1 Introduction 240
11.2 The San Diego SWIM Program Data 240
11.3 Fisher's Exact P-Values 242
11.4 Neyman's Repeated Sampling-Based Point Estimates and Large-Sample Confidence Intervals 245
11.5 Regression-Based Estimates 247
11.6 Model-Based Point Estimates 250
11.7 Conclusion 253
Notes 253

PART III REGULAR ASSIGNMENT MECHANISMS: DESIGN

12 Unconfounded Treatment Assignment 257
12.1 Introduction 257
12.2 Regular Assignment Mechanisms 258
12.3 Balancing Scores and the Propensity Score 266
12.4 Estimation and Inference 268
12.5 Design Phase 276

12.6 Assessing Unconfoundedness 278
12.7 Conclusion 279
Notes 279

13 Estimating the Propensity Score 281
13.1 Introduction 281
13.2 The Reinisch et al. Barbituate Exposure Data 284
13.3 Selecting the Covariates and Interactions 285
13.4 Choosing the Specification of the Propensity Score for the Barbituate Data 288
13.5 Constructing Propensity-Score Strata 290
13.6 Choosing Strata for the Barbituate Data 294
13.7 Assessing Balance Conditional on the Estimated Propensity Score 296
13.8 Assessing Covariate Balance for the Barbituate Data 300
13.9 Conclusion 306
Notes 306
Appendix: Logistic Regression 307

14 Assessing Overlap in Covariate Distributions 309
14.1 Introduction 309
14.2 Assessing Balance in Univariate Distributions 310
14.3 Direct Assessment of Balance in Multivariate Distributions 313
14.4 Assessing Balance in Multivariate Distributions Using the Propensity Score 314
14.5 Assessing the Ability to Adjust for Differences in Covariates by Treatment Status 317
14.6 Assessing Balance: Four Illustrations 318
14.7 Sensitivity of Regression Estimates to Lack of Overlap 332
14.8 Conclusion 336
Notes 336

15 Matching to Improve Balance in Covariate Distributions 337
15.1 Introduction 337
15.2 The Reinisch et al. Barbituate Exposure Data 339
15.3 Selecting a Subsample of Controls through Matching to Improve Balance 339
15.4 An Illustration of Propensity Score Matching with Six Observations 344
15.5 Theoretical Properties of Matching Procedures 345
15.6 Creating Matched Samples for the Barbituate Data 349
15.7 Conclusion 358
Notes 358

16 Trimming to Improve Balance in Covariate Distributions 359
16.1 Introduction 359
16.2 The Right Heart Catheterization Data 360
16.3 An Example with a Single Binary Covariate 362

16.4 Selecting a Subsample Based on the Propensity Score 366
16.5 The Optimal Subsample for the Right Heart Catheterization Data 368
16.6 Conclusion 373
Notes 374

PART IV REGULAR ASSIGNMENT MECHANISMS: ANALYSIS

17 Subclassification on the Propensity Score 377
17.1 Introduction 377
17.2 The Imbens-Rubin-Sacerdote Lottery Data 378
17.3 Subclassification on the Propensity Score and Bias Reduction 380
17.4 Subclassification and the Lottery Data 385
17.5 Estimation Based on Subclassification with Additional Bias Reduction 386
17.6 Neymanian Inference 388
17.7 Average Treatment Effects for the Lottery Data 390
17.8 Weighting Estimators and Subclassification 392
17.9 Conclusion 399
Notes 399

18 Matching Estimators 401
18.1 Introduction 401
18.2 The Card-Krueger New Jersey and Pennsylvania Minimum Wage Data 404
18.3 Exact Matching without Replacement 405
18.4 Inexact Matching without Replacement 407
18.5 Distance Measures 410
18.6 Matching and the Card-Krueger Data 412
18.7 The Bias of Matching Estimators 415
18.8 Bias-Corrected Matching Estimators 416
18.9 Matching with Replacement 424
18.10 The Number of Matches 425
18.11 Matching Estimators for the Average Treatment Effect for the Controls and for the Full Sample 427
18.12 Matching Estimates of the Effect of the Minimum Wage Increase 428
18.13 Conclusion 430
Notes 431

19 A General Method for Estimating Sampling Variances for Standard Estimators for Average Causal Effects 433
19.1 Introduction 433
19.2 The Imbens-Rubin-Sacerdote Lottery Data 435
19.3 Estimands 436

19.4 The Common Structure of Standard Estimators for Average Treatment Effects 441
19.5 A General Formula for the Conditional Sampling Variance 445
19.6 A Simple Estimator for the Unit-Level Conditional Sampling Variance 446
19.7 An Estimator for the Sampling Variance of $\hat{\tau}$ Conditional on Covariates 452
19.8 An Estimator for the Sampling Variance for the Estimator for the Average Effect for the Treated 452
19.9 An Estimator for the Sampling Variance for the Population Average Treatment Effect 454
19.10 Alternative Estimators for the Sampling Variance 456
19.11 Conclusion 460
Notes 460

20 Inference for General Causal Estimands 461
20.1 Introduction 461
20.2 The Lalonde NSW Observational Job-Training Data 462
20.3 Causal Estimands 465
20.4 A Model for the Conditional Potential Outcome Distributions 468
20.5 Implementation 472
20.6 Results for the Lalonde Data 473
20.7 Conclusion 474
Notes 474

PART V REGULAR ASSIGNMENT MECHANISMS: SUPPLEMENTARY ANALYSES

21 Assessing Unconfoundedness 479
21.1 Introduction 479
21.2 Setup 482
21.3 Estimating Effects on Pseudo-Outcomes 482
21.4 Estimating Effects of Pseudo-Treatments 485
21.5 Robustness to the Set of Pre-Treatment Variables 487
21.6 The Imbens-Rubin-Sacerdote Lottery Data 490
21.7 Conclusion 495
Notes 495

22 Sensitivity Analysis and Bounds 496
22.1 Introduction 496
22.2 The Imbens-Rubin-Sacerdote Lottery Data 497
22.3 Bounds 497
22.4 Binary Outcomes: The Rosenbaum-Rubin Sensitivity Analysis 500

22.5 Binary Outcomes: The Rosenbaum Sensitivity Analysis for P-Values 506
22.6 Conclusion 509
Notes 509

PART VI REGULAR ASSIGNMENT MECHANISMS WITH NONCOMPLIANCE: ANALYSIS

23 Instrumental Variables Analysis of Randomized Experiments with One-Sided Noncompliance 513
23.1 Introduction 513
23.2 The Sommer-Zeger Vitamin A Supplement Data 516
23.3 Setup 517
23.4 Intention-to-Treat Effects 519
23.5 Compliance Status 522
23.6 Instrumental Variables 526
23.7 Moment-Based Instrumental Variables Estimators 530
23.8 Linear Models and Instrumental Variables 531
23.9 Naive Analyses: "As-Treated," "Per Protocol," and Unconfoundedness 535
23.10 Conclusion 539
Notes 539
Appendix 541

24 Instrumental Variables Analysis of Randomized Experiments with Two-Sided Noncompliance 542
24.1 Introduction 542
24.2 The Angrist Draft Lottery Data 543
24.3 Compliance Status 544
24.4 Intention-to-Treat Effects 546
24.5 Instrumental Variables 548
24.6 Traditional Econometric Methods for Instrumental Variables 556
24.7 Conclusion 559
Notes 559

25 Model-Based Analysis in Instrumental Variable Settings: Randomized Experiments with Two-Sided Noncompliance 560
25.1 Introduction 560
25.2 The McDonald-Hiu-Tierney Influenza Vaccination Data 561
25.3 Covariates 567
25.4 Model-Based Instrumental Variables Analyses for Randomized Experiments with Two-Sided Noncompliance 568
25.5 Simulation Methods for Obtaining Draws from the Posterior Distribution of the Estimand Given the Data 574
25.6 Models for the Influenza Vaccination Data 578

25.7 Results for the Influenza Vaccination Data 581
25.8 Conclusion 584
Notes 584

PART VII CONCLUSION

26 Conclusions and Extensions 589
Notes 590

References 591
Author Index 605
Subject Index 609