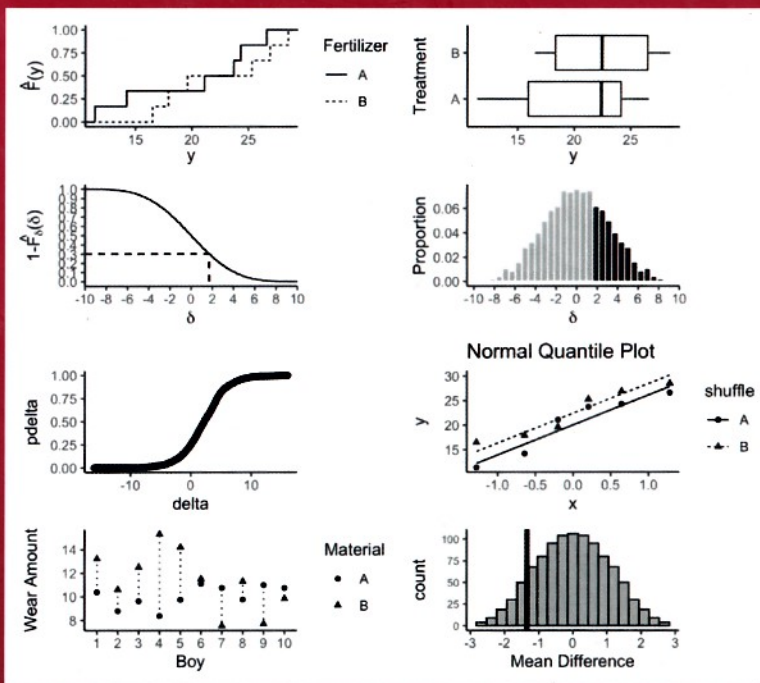


Texts in Statistical Science

# Design and Analysis of Experiments and Observational Studies using R



Nathan Taback



CRC Press  
Taylor & Francis Group

A CHAPMAN & HALL BOOK

---

# *Contents*

---

List of Tables	xiii
List of Figures	xv
Symbols and Abbreviations	xvii
Preface	xix
<b>1 Introduction</b>	<b>1</b>
1.1 Planning and Implementing a Study . . . . .	1
1.2 Questions . . . . .	2
1.3 Types of Study Investigations . . . . .	3
1.4 Strength of Answers . . . . .	4
1.5 Why Design Scientific Studies? . . . . .	4
<b>2 Mathematical Statistics: Simulation and Computation</b>	<b>7</b>
2.1 Data . . . . .	7
2.1.1 Computation Lab: Data . . . . .	7
2.2 Frequency Distributions . . . . .	13
2.2.1 Discrete and Continuous Data . . . . .	13
2.2.2 Computation Lab: Frequency Distributions . . . . .	14
2.3 Randomization . . . . .	18
2.3.1 Computation Lab: Randomization . . . . .	19
2.4 Theoretical Distributions . . . . .	21
2.4.1 The Binomial Distribution . . . . .	22
2.4.2 The Normal Distribution . . . . .	22
2.4.3 Chi-Square Distribution . . . . .	24
2.4.4 t Distribution . . . . .	24
2.4.5 F Distribution . . . . .	26
2.4.6 Computation Lab: Theoretical Distributions . . . . .	26
2.5 Quantile-Quantile Plots . . . . .	29
2.5.1 Computation Lab: Quantile-Quantile Plots . . . . .	31
2.6 Central Limit Theorem . . . . .	34
2.6.1 Computation Lab: Central Limit Theorem . . . . .	34
2.7 Statistical Inference . . . . .	37
2.7.1 Hypothesis Testing . . . . .	37
2.7.2 Hypothesis Testing via Simulation . . . . .	37

2.7.3	Model-based Hypothesis Testing . . . . .	38
2.7.4	Computation Lab: Hypothesis Testing . . . . .	38
2.7.5	Confidence Intervals . . . . .	41
2.7.6	Computation Lab: Confidence Intervals . . . . .	42
2.7.7	Errors in Hypothesis Testing . . . . .	43
2.8	Linear Regression . . . . .	44
2.8.1	Simple Linear Regression . . . . .	44
2.8.2	Multiple Linear Regression . . . . .	45
2.8.3	Categorical Covariates in Regression . . . . .	46
2.8.4	Computation Lab: Linear Regression . . . . .	49
2.9	Exercises . . . . .	52
<b>3</b>	<b>Comparing Two Treatments</b>	<b>57</b>
3.1	Introduction . . . . .	57
3.2	Treatment Assignment Mechanism and Propensity Score . . . . .	58
3.2.1	Propensity Score . . . . .	58
3.2.2	Assignment Mechanism . . . . .	58
3.2.3	Computation Lab: Treatment Assignment Mechanism and Propensity Score . . . . .	59
3.3	Completely Randomized Designs . . . . .	59
3.3.1	Computation Lab: Completely Randomized Experiments . . . . .	60
3.4	The Randomization Distribution . . . . .	61
3.4.1	Computation Lab: Randomization Distribution . . . . .	62
3.5	The Randomization p-value . . . . .	63
3.5.1	One-sided p-value . . . . .	63
3.5.2	Two-sided Randomization p-value . . . . .	64
3.5.3	Computation Lab: Randomization p-value . . . . .	64
3.5.4	Randomization Confidence Intervals . . . . .	66
3.5.5	Computation Lab: Randomization Confidence Intervals . . . . .	67
3.6	Randomization Distribution of a Test Statistic . . . . .	70
3.6.1	Computation Lab: Randomization Distribution of a Test Statistic . . . . .	70
3.7	Computing the Randomization Distribution using Monte Carlo Sampling . . . . .	71
3.7.1	Computation Lab: Calculating the Randomization Distribution using Monte Carlo Sampling . . . . .	71
3.8	Properties of the Randomization Test . . . . .	73
3.9	The Two-sample t-test . . . . .	73
3.9.1	Computation Lab: Two-sample t-test . . . . .	74
3.10	Blocking . . . . .	77
3.11	Treatment Assignment in Randomized Block Designs . . . . .	78
3.11.1	Computation Lab: Generating a Randomized Block Design . . . . .	78
3.12	Randomized Matched Pairs Design . . . . .	79

3.13	Randomized Matched Pairs versus Completely Randomized Design . . . . .	80
3.14	The Randomization Test for a Randomized Paires Design . . . . .	81
3.14.1	Computation Lab: Randomization Test for a Paired Design . . . . .	82
3.15	Paired t-test . . . . .	84
3.15.1	Computation Lab: Paired t-test . . . . .	84
3.16	Exercises . . . . .	85
<b>4</b>	<b>Power and Sample Size</b>	<b>89</b>
4.1	Introduction . . . . .	89
4.2	Statistical Hypotheses and the Number of Experimental Units	90
4.3	Power of the One-Sample z-test . . . . .	91
4.3.1	Computation Lab: Power of the One-Sample z-test . . . . .	92
4.4	Power of the One-Sample t-test . . . . .	92
4.4.1	Computation Lab: Power of the One-Sample t-test . . . . .	94
4.5	Power of the Two-Sample t-test . . . . .	95
4.5.1	Effect Size . . . . .	96
4.5.2	Sample Size—Known Variance and Equal Allocation . . . . .	96
4.5.3	Sample Size—Known Variance and Unequal Allocation . . . . .	97
4.5.4	Computation Lab: Power of the Two-Sample t-test . . . . .	98
4.6	Power and Sample Size for Comparing Proportions . . . . .	105
4.6.1	Computation Lab: Power and Sample Size for Comparing Proportions . . . . .	106
4.7	Calculating Power by Simulation . . . . .	107
4.7.1	Algorithm for Simulating Power . . . . .	107
4.7.2	Computation Lab: Simulating Power of a Two-Sample t-test . . . . .	107
4.8	Exercises . . . . .	111
<b>5</b>	<b>Comparing More Than Two Treatments</b>	<b>115</b>
5.1	Introduction: ANOVA—Comparing More Than Two Groups	115
5.2	Random Assignment of Treatments . . . . .	116
5.2.1	Computation Lab: Random Assignment . . . . .	116
5.3	ANOVA . . . . .	117
5.3.1	Source of Variation and Sums of Squares . . . . .	117
5.3.2	ANOVA Identity . . . . .	118
5.3.3	Computation Lab: ANOVA Identity . . . . .	119
5.3.4	ANOVA: Degrees of Freedom . . . . .	120
5.3.5	ANOVA: Mean Squares . . . . .	120
5.3.6	ANOVA: F Statistic . . . . .	121
5.3.7	ANOVA Table . . . . .	121
5.3.8	Computation Lab: ANOVA Table . . . . .	121
5.4	Estimating Treatment Effects Using Least Squares . . . . .	122
5.4.1	Dummy Coding . . . . .	122
5.4.2	Deviation Coding . . . . .	123

5.5	Computation Lab: Estimating Treatment Effects Using Least Squares . . . . .	124
5.5.1	ANOVA Assumptions . . . . .	127
5.5.2	Computation Lab: Checking ANOVA Assumptions . . . . .	127
5.6	Multiple Comparisons . . . . .	128
5.6.1	The Bonferroni Method . . . . .	130
5.6.2	The Tukey Method . . . . .	130
5.6.3	Computation Lab: Multiple Comparisons . . . . .	131
5.7	Sample Size for ANOVA—Designing a Study to Compare More Than Two Treatments . . . . .	132
5.7.1	Computation Lab: Sample Size for ANOVA . . . . .	134
5.8	Randomized Block Designs . . . . .	138
5.8.1	ANOVA Identity for Randomized Block Designs . . . . .	138
5.8.2	Computation Lab: ANOVA for Randomized Block Design . . . . .	140
5.9	The Linear Model for Randomized Block Design . . . . .	143
5.9.1	Computation Lab: Regression Model for Randomized Block Design . . . . .	144
5.9.2	Computation Lab: Balanced Randomization via Permuted Block Randomization . . . . .	148
5.10	Latin Square Designs . . . . .	149
5.10.1	Computation Lab: Randomizing Treatments to a Latin Square . . . . .	151
5.11	Statistical Analysis of Latin Square Design . . . . .	153
5.11.1	Computation Lab: Analysis of Latin Square Designs . . . . .	154
5.12	Graeco-Latin Square Designs . . . . .	160
5.12.1	Computation Lab: Graeco-Latin Square Designs . . . . .	161
5.13	Exercises . . . . .	161
<b>6</b>	<b>Factorial Designs at Two Levels—<math>2^k</math> Designs</b>	<b>167</b>
6.1	Introduction . . . . .	167
6.2	Factorial Effects . . . . .	168
6.2.1	Main Effects . . . . .	168
6.2.2	Interaction Effects . . . . .	169
6.2.3	Three Factor Interactions . . . . .	170
6.2.4	Cube Plot and Model Matrix . . . . .	171
6.2.5	Computation Lab: Computing and Plotting Factorial Effects . . . . .	172
6.3	Replication in Factorial Designs . . . . .	178
6.3.1	Computation Lab: Replication in Factorial Designs . . . . .	182
6.4	Linear Model for a $2^k$ Factorial Design . . . . .	185
6.4.1	Computation Lab: Linear Model for a $2^k$ Factorial Design . . . . .	186
6.5	Normal Plots in Unreplicated Factorial Designs . . . . .	190
6.5.1	Half-Normal Plots . . . . .	192
6.5.2	Computation Lab: Normal Plots in Unreplicated Factorial Designs . . . . .	193

6.6	Lenth's Method . . . . .	195
6.6.1	Computation Lab: Lenth's method . . . . .	196
6.7	Blocking Factorial Designs . . . . .	199
6.7.1	Effect Heirarchy Principle . . . . .	200
6.7.2	Generation of Orthogonal Blocks . . . . .	201
6.7.3	Generators and Defining Relations . . . . .	202
6.7.4	Computation Lab: Blocking Factorial Designs . . . . .	203
6.8	Fractional Factorial Designs . . . . .	205
6.8.1	Effect Aliasing and Design Resolution . . . . .	207
6.8.2	Computation Lab: Fractional Factorial Designs . . . . .	210
6.9	Exercises . . . . .	212
<b>7</b>	<b>Causal Inference</b>	<b>219</b>
7.1	Introduction: The Fundamental Problem of Causal Inference	219
7.1.1	Stable Unit Treatment Value Assumption (SUTVA)	221
7.2	Treatment Assignment . . . . .	222
7.2.1	Computation Lab: Treatment Assignment . . . . .	226
7.3	Causal Effects and Randomized Experiments . . . . .	228
7.4	Causal Effects and Observational Studies . . . . .	230
7.4.1	What is an Observational Study? . . . . .	230
7.4.2	Designing an Observational Study . . . . .	231
7.5	Epidemiologic Follow-up Study . . . . .	231
7.6	The Propensity Score . . . . .	233
7.6.1	The Balancing Property of the Propensity Score . . . . .	234
7.6.2	Propensity Score and Ignorable Treatment Assignment	235
7.7	Propensity Score Methods to Reduce Bias in Observational Studies . . . . .	236
7.7.1	Propensity Score Matching . . . . .	237
7.7.2	Distance Between Units . . . . .	237
7.7.3	Nearest Neighbour Matching . . . . .	238
7.7.4	Computation Lab: Nearest Neighbour Propensity Score Matching . . . . .	238
7.7.5	Assessing Covariate Balance in the Matched Sample . . . . .	241
7.7.6	Computation Lab: Checking Overlap and Covariate Balance . . . . .	243
7.7.7	Analysis of the Outcome after Matching . . . . .	246
7.7.8	Computation Lab: Analysis of Outcome after Matching	247
7.7.9	Subclassification . . . . .	249
7.7.10	Computation Lab: Subclassification and the Propensity Score . . . . .	251
7.8	Exercises . . . . .	256

<b>Bibliography</b>	<b>261</b>
---------------------	------------

<b>Index</b>	<b>267</b>
--------------	------------

***Introduction to Design and Analysis of Scientific Studies*** exposes undergraduate and graduate students to the foundations of classical experimental design and observational studies through a modern framework - The Rubin Causal Model. A causal inference framework is important in design, data collection and analysis since it provides a framework for investigators to readily evaluate study limitations and draw appropriate conclusions. R is used to implement designs and analyse the data collected.

## Features


- Classical experimental design with an emphasis on computation using tidyverse packages in R.
- Applications of experimental design to clinical trials, A/B testing, and other modern examples.
- Discussion of the link between classical experimental design and causal inference.
- The role of randomization in experimental design and sampling in the big data era.
- Exercises with solutions.

Instructor slides in RMarkdown, a new R package will be developed to be used with book, and a bookdown version of the book will be freely available. The proposed book will emphasize ethics, communication and decision making as part of design, data analysis, and statistical thinking.

## About the Author

**Nathan Taback** is Associate Professor of Statistics and Data Science at University of Toronto.

STATISTICS

 **CRC Press**  
Taylor & Francis Group  
an **informa** business  
[www.routledge.com](http://www.routledge.com)

CRC Press titles are available as eBook editions in a range of digital formats

