

# Contents

<b>Introduction: Towards less causal causal inferences</b>	<b>vii</b>
<b>I Causal inference without models</b>	<b>1</b>
<b>1 A definition of causal effect</b>	<b>3</b>
1.1 Individual causal effects . . . . .	3
1.2 Average causal effects . . . . .	4
1.3 Measures of causal effect . . . . .	7
1.4 Random variability . . . . .	8
1.5 Causation versus association . . . . .	10
<b>2 Randomized experiments</b>	<b>13</b>
2.1 Randomization . . . . .	13
2.2 Conditional randomization . . . . .	17
2.3 Standardization . . . . .	19
2.4 Inverse probability weighting . . . . .	20
<b>3 Observational studies</b>	<b>27</b>
3.1 Identifiability conditions . . . . .	27
3.2 Exchangeability . . . . .	29
3.3 Positivity . . . . .	32
3.4 Consistency: First, define the counterfactual outcome . . . . .	33
3.5 Consistency: Second, link counterfactuals to the observed data .	37
3.6 The target trial . . . . .	38
<b>4 Effect modification</b>	<b>43</b>
4.1 Heterogeneity of treatment effects . . . . .	43
4.2 Stratification to identify effect modification . . . . .	45
4.3 Why care about effect modification . . . . .	47
4.4 Stratification as a form of adjustment . . . . .	49
4.5 Matching as another form of adjustment . . . . .	51
4.6 Effect modification and adjustment methods . . . . .	52
<b>5 Interaction</b>	<b>57</b>
5.1 Interaction requires a joint intervention . . . . .	57
5.2 Identifying interaction . . . . .	58
5.3 Counterfactual response types and interaction . . . . .	60
5.4 Sufficient causes . . . . .	62
5.5 Sufficient cause interaction . . . . .	65
5.6 Counterfactuals or sufficient-component causes? . . . . .	67

<b>6 Graphical representation of causal effects</b>	<b>71</b>
6.1 Causal diagrams . . . . .	71
6.2 Causal diagrams and marginal independence . . . . .	73
6.3 Causal diagrams and conditional independence . . . . .	76
6.4 Positivity and consistency in causal diagrams . . . . .	77
6.5 A structural classification of bias . . . . .	80
6.6 The structure of effect modification . . . . .	83
<b>7 Confounding</b>	<b>85</b>
7.1 The structure of confounding . . . . .	85
7.2 Confounding and exchangeability . . . . .	87
7.3 Confounding and the backdoor criterion . . . . .	89
7.4 Confounding and confounders . . . . .	92
7.5 Single-world intervention graphs . . . . .	95
7.6 Confounding adjustment . . . . .	96
<b>8 Selection bias</b>	<b>103</b>
8.1 The structure of selection bias . . . . .	103
8.2 Examples of selection bias . . . . .	105
8.3 Selection bias and confounding . . . . .	107
8.4 Selection bias and censoring . . . . .	109
8.5 How to adjust for selection bias . . . . .	111
8.6 Selection without bias . . . . .	115
<b>9 Measurement bias and “Noncausal” diagrams</b>	<b>119</b>
9.1 Measurement error . . . . .	119
9.2 The structure of measurement error . . . . .	120
9.3 Mismeasured confounders and colliders . . . . .	122
9.4 Causal diagrams without mismeasured variables? . . . . .	124
9.5 Many proposed causal diagrams include noncausal arrows . . . . .	125
9.6 Does it matter that many proposed diagrams include noncausal arrows? . . . . .	128
<b>10 Random variability</b>	<b>131</b>
10.1 Identification versus estimation . . . . .	131
10.2 Estimation of causal effects . . . . .	134
10.3 The myth of the super-population . . . . .	136
10.4 The conditionality “principle” . . . . .	138
10.5 The curse of dimensionality . . . . .	142
<b>II Causal inference with models</b>	<b>145</b>
<b>11 Why model?</b>	<b>147</b>
11.1 Data cannot speak for themselves . . . . .	147
11.2 Parametric estimators of the conditional mean . . . . .	149
11.3 Nonparametric estimators of the conditional mean . . . . .	150
11.4 Smoothing . . . . .	151
11.5 The bias-variance trade-off . . . . .	153

<b>12 IP weighting and marginal structural models</b>	<b>157</b>
12.1 The causal question . . . . .	157
12.2 Estimating IP weights via modeling . . . . .	158
12.3 Stabilized IP weights . . . . .	161
12.4 Marginal structural models . . . . .	163
12.5 Effect modification and marginal structural models . . . . .	165
12.6 Censoring and missing data . . . . .	166
<b>13 Standardization and the parametric g-formula</b>	<b>169</b>
13.1 Standardization as an alternative to IP weighting . . . . .	169
13.2 Estimating the mean outcome via modeling . . . . .	171
13.3 Standardizing the mean outcome to the confounder distribution	172
13.4 IP weighting or standardization? . . . . .	173
13.5 How seriously do we take our estimates? . . . . .	175
<b>14 G-estimation of structural nested models</b>	<b>181</b>
14.1 The causal question revisited . . . . .	181
14.2 Exchangeability revisited . . . . .	182
14.3 Structural nested mean models . . . . .	183
14.4 Rank preservation . . . . .	185
14.5 G-estimation . . . . .	187
14.6 Structural nested models with two or more parameters . . . . .	190
<b>15 Outcome regression and propensity scores</b>	<b>193</b>
15.1 Outcome regression . . . . .	193
15.2 Propensity scores . . . . .	195
15.3 Propensity stratification and standardization . . . . .	196
15.4 Propensity matching . . . . .	198
15.5 Propensity models, structural models, predictive models . . . . .	199
<b>16 Instrumental variable estimation</b>	<b>203</b>
16.1 The three instrumental conditions . . . . .	203
16.2 The usual IV estimand . . . . .	206
16.3 A fourth identifying condition: homogeneity . . . . .	208
16.4 An alternative fourth condition: monotonicity . . . . .	211
16.5 The three instrumental conditions revisited . . . . .	214
16.6 Instrumental variable estimation versus other methods . . . . .	217
<b>17 Causal survival analysis</b>	<b>221</b>
17.1 Hazards and risks . . . . .	221
17.2 From hazards to risks . . . . .	223
17.3 Why censoring matters . . . . .	226
17.4 IP weighting of marginal structural models . . . . .	228
17.5 The parametric g-formula . . . . .	230
17.6 G-estimation of structural nested models . . . . .	231
<b>18 Variable selection and high-dimensional data</b>	<b>235</b>
18.1 The different goals of variable selection . . . . .	235
18.2 Variables that induce or amplify bias . . . . .	236
18.3 Causal inference and machine learning . . . . .	240
18.4 Doubly robust machine learning estimators . . . . .	241
18.5 Variable selection is a difficult problem . . . . .	244

**III Causal inference for time-varying treatments 247**

<b>19 Time-varying treatments</b>	<b>249</b>
19.1 The causal effect of time-varying treatments . . . . .	249
19.2 Treatment strategies . . . . .	250
19.3 Sequentially randomized experiments . . . . .	251
19.4 Sequential exchangeability . . . . .	253
19.5 Identifiability under some but not all treatment strategies . . . .	255
19.6 Time-varying confounding and time-varying confounders . . . .	259
<b>20 Treatment-confounder feedback</b>	<b>261</b>
20.1 The elements of treatment-confounder feedback . . . . .	261
20.2 The bias of traditional methods . . . . .	263
20.3 Why traditional methods fail . . . . .	265
20.4 Why traditional methods cannot be fixed . . . . .	267
20.5 Adjusting for past treatment . . . . .	268
<b>21 G-methods for time-varying treatments</b>	<b>271</b>
21.1 The g-formula for time-varying treatments . . . . .	271
21.2 IP weighting for time-varying treatments . . . . .	276
21.3 A doubly robust estimator for time-varying treatments . . . .	280
21.4 G-estimation for time-varying treatments . . . . .	283
21.5 Censoring is a time-varying treatment . . . . .	291
21.6 The big g-formula . . . . .	294
<b>22 Target trial emulation</b>	<b>299</b>
22.1 Intention-to-treat effect and per-protocol effect . . . . .	299
22.2 A target trial with sustained treatment strategies . . . . .	303
22.3 Emulating a target trial with sustained strategies . . . .	307
22.4 Time zero . . . . .	309
22.5 A unified approach to answer What If questions with data . . .	311
<b>23 Causal mediation</b>	<b>317</b>
23.1 Mediation analysis under attack . . . . .	317
23.2 A defense of mediation analysis . . . . .	319
23.3 Empirically verifiable mediation . . . . .	321
23.4 An interventionist theory of mediation . . . . .	323
<b>References</b>	<b>326</b>
<b>Index</b>	<b>345</b>