

CAUSALITY

Models, Reasoning, and Inference

Judea Pearl

University of California, Los Angeles



CAMBRIDGE
UNIVERSITY PRESS

Contents

<i>Preface</i>	<i>page</i>	<i>xiii</i>
1 Introduction to Probabilities, Graphs, and Causal Models		1
1.1 Introduction to Probability Theory		1
1.1.1 Why Probabilities?		1
1.1.2 Basic Concepts in Probability Theory		2
1.1.3 Combining Predictive and Diagnostic Supports		6
1.1.4 Random Variables and Expectations		8
1.1.5 Conditional Independence and Graphoids		11
1.2 Graphs and Probabilities		12
1.2.1 Graphical Notation and Terminology		12
1.2.2 Bayesian Networks		13
1.2.3 The d -Separation Criterion		16
1.2.4 Inference with Bayesian Networks		20
1.3 Causal Bayesian Networks		21
1.3.1 Causal Networks as Oracles for Interventions		22
1.3.2 Causal Relationships and Their Stability		24
1.4 Functional Causal Models		26
1.4.1 Structural Equations		27
1.4.2 Probabilistic Predictions in Causal Models		30
1.4.3 Interventions and Causal Effects in Functional Models		32
1.4.4 Counterfactuals in Functional Models		33
1.5 Causal versus Statistical Terminology		38
2 A Theory of Inferred Causation		41
2.1 Introduction		42
2.2 The Causal Modeling Framework		43
2.3 Model Preference (Occam's Razor)		45
2.4 Stable Distributions		48
2.5 Recovering DAG Structures		49
2.6 Recovering Latent Structures		51
2.7 Local Criteria for Causal Relations		54

2.8	Nontemporal Causation and Statistical Time	57
2.9	Conclusions	59
2.9.1	On Minimality, Markov, and Stability	61
3	Causal Diagrams and the Identification of Causal Effects	65
3.1	Introduction	66
3.2	Intervention in Markovian Models	68
3.2.1	Graphs as Models of Interventions	68
3.2.2	Interventions as Variables	70
3.2.3	Computing the Effect of Interventions	72
3.2.4	Identification of Causal Quantities	77
3.3	Controlling Confounding Bias	78
3.3.1	The Back-Door Criterion	79
3.3.2	The Front-Door Criterion	81
3.3.3	Example: Smoking and the Genotype Theory	83
3.4	A Calculus of Intervention	85
3.4.1	Preliminary Notation	85
3.4.2	Inference Rules	85
3.4.3	Symbolic Derivation of Causal Effects: An Example	86
3.4.4	Causal Inference by Surrogate Experiments	88
3.5	Graphical Tests of Identifiability	89
3.5.1	Identifying Models	91
3.5.2	Nonidentifying Models	93
3.6	Discussion	94
3.6.1	Qualifications and Extensions	94
3.6.2	Diagrams as a Mathematical Language	96
3.6.3	Translation from Graphs to Potential Outcomes	98
3.6.4	Relations to Robins's <i>G</i> -Estimation	102
4	Actions, Plans, and Direct Effects	107
4.1	Introduction	108
4.1.1	Actions, Acts, and Probabilities	108
4.1.2	Actions in Decision Analysis	110
4.1.3	Actions and Counterfactuals	112
4.2	Conditional Actions and Stochastic Policies	113
4.3	When Is the Effect of an Action Identifiable?	114
4.3.1	Graphical Conditions for Identification	114
4.3.2	Remarks on Efficiency	116
4.3.3	Deriving a Closed-Form Expression for Control Queries	117
4.3.4	Summary	118
4.4	The Identification of Plans	118
4.4.1	Motivation	118
4.4.2	Plan Identification: Notation and Assumptions	120
4.4.3	Plan Identification: A General Criterion	121
4.4.4	Plan Identification: A Procedure	124

4.5	Direct Effects and Their Identification	126
4.5.1	Direct versus Total Effects	126
4.5.2	Direct Effects, Definition, and Identification	127
4.5.3	Example: Sex Discrimination in College Admission	128
4.5.4	Average Direct Effects	130
5	Causality and Structural Models in Social Science and Economics	133
5.1	Introduction	134
5.1.1	Causality in Search of a Language	134
5.1.2	SEM: How its Meaning Became Obscured	135
5.1.3	Graphs as a Mathematical Language	138
5.2	Graphs and Model Testing	140
5.2.1	The Testable Implications of Structural Models	140
5.2.2	Testing the Testable	144
5.2.3	Model Equivalence	145
5.3	Graphs and Identifiability	149
5.3.1	Parameter Identification in Linear Models	149
5.3.2	Comparison to Nonparametric Identification	154
5.3.3	Causal Effects: The Interventional Interpretation of Structural Equation Models	157
5.4	Some Conceptual Underpinnings	159
5.4.1	What Do Structural Parameters Really Mean?	159
5.4.2	Interpretation of Effect Decomposition	163
5.4.3	Exogeneity, Superexogeneity, and Other Frills	165
5.5	Conclusion	170
6	Simpson's Paradox, Confounding, and Collapsibility	173
6.1	Simpson's Paradox: An Anatomy	174
6.1.1	A Tale of a Non-Paradox	174
6.1.2	A Tale of Statistical Agony	175
6.1.3	Causality versus Exchangeability	177
6.1.4	A Paradox Resolved (Or: What Kind of Machine Is Man?)	180
6.2	Why There Is No Statistical Test for Confounding, Why Many Think There Is, and Why They Are Almost Right	182
6.2.1	Introduction	182
6.2.2	Causal and Associational Definitions	184
6.3	How the Associational Criterion Fails	185
6.3.1	Failing Sufficiency via Marginality	185
6.3.2	Failing Sufficiency via Closed-World Assumptions	186
6.3.3	Failing Necessity via Barren Proxies	186
6.3.4	Failing Necessity via Incidental Cancellations	188
6.4	Stable versus Incidental Unbiasedness	189
6.4.1	Motivation	189
6.4.2	Formal Definitions	191
6.4.3	Operational Test for Stable No-Confounding	192

6.5	Confounding, Collapsibility, and Exchangeability	193
6.5.1	Confounding and Collapsibility	193
6.5.2	Counfounding versus Confounders	194
6.5.3	Exchangeability versus Structural Analysis of Confounding	196
6.6	Conclusions	199
7	The Logic of Structure-Based Counterfactuals	201
7.1	Structural Model Semantics	202
7.1.1	Definitions: Causal Models, Actions, and Counterfactuals	202
7.1.2	Evaluating Counterfactuals: Deterministic Analysis	207
7.1.3	Evaluating Counterfactuals: Probabilistic Analysis	212
7.1.4	The Twin Network Method	213
7.2	Applications and Interpretation of Structural Models	215
7.2.1	Policy Analysis in Linear Econometric Models: An Example	215
7.2.2	The Empirical Content of Counterfactuals	217
7.2.3	Causal Explanations, Utterances, and Their Interpretation	221
7.2.4	From Mechanisms to Actions to Causation	223
7.2.5	Simon's Causal Ordering	226
7.3	Axiomatic Characterization	228
7.3.1	The Axioms of Structural Counterfactuals	228
7.3.2	Causal Effects from Counterfactual Logic: An Example	231
7.3.3	Axioms of Causal Relevance	234
7.4	Structural and Similarity-Based Counterfactuals	238
7.4.1	Relations to Lewis's Counterfactuals	238
7.4.2	Axiomatic Comparison	240
7.4.3	Imaging versus Conditioning	242
7.4.4	Relations to the Neyman–Rubin Framework	243
7.4.5	Exogeneity Revisited: Counterfactual and Graphical Definitions	245
7.5	Structural versus Probabilistic Causality	249
7.5.1	The Reliance on Temporal Ordering	249
7.5.2	The Perils of Circularity	250
7.5.3	The Closed-World Assumption	252
7.5.4	Singular versus General Causes	253
7.5.5	Summary	256
8	Imperfect Experiments: Bounding Effects and Counterfactuals	259
8.1	Introduction	259
8.1.1	Imperfect and Indirect Experiments	259
8.1.2	Noncompliance and Intent to Treat	261
8.2	Bounding Causal Effects	262
8.2.1	Problem Formulation	262
8.2.2	The Evolution of Potential-Response Variables	263
8.2.3	Linear Programming Formulation	266

8.2.4	The Natural Bounds	268
8.2.5	Effect of Treatment on the Treated	269
8.2.6	Example: The Effect of Cholestyramine	270
8.3	Counterfactuals and Legal Responsibility	271
8.4	A Test for Instruments	274
8.5	Causal Inference from Finite Samples	275
8.5.1	Gibbs Sampling	275
8.5.2	The Effects of Sample Size and Prior Distribution	277
8.5.3	Causal Effects from Clinical Data with Imperfect Compliance	277
8.5.4	Bayesian Estimate of Single-Event Causation	280
8.6	Conclusion	281
9	Probability of Causation: Interpretation and Identification	283
9.1	Introduction	283
9.2	Necessary and Sufficient Causes: Conditions of Identification	286
9.2.1	Definitions, Notation, and Basic Relationships	286
9.2.2	Bounds and Basic Relationships under Exogeneity	289
9.2.3	Identifiability under Monotonicity and Exogeneity	291
9.2.4	Identifiability under Monotonicity and Nonexogeneity	293
9.3	Examples and Applications	296
9.3.1	Example 1: Betting against a Fair Coin	297
9.3.2	Example 2: The Firing Squad	297
9.3.3	Example 3: The Effect of Radiation on Leukemia	299
9.3.4	Example 4: Legal Responsibility from Experimental and Nonexperimental Data	302
9.3.5	Summary of Results	303
9.4	Identification in Nonmonotonic Models	304
9.5	Conclusions	307
10	The Actual Cause	309
10.1	Introduction: The Insufficiency of Necessary Causation	309
10.1.1	Singular Causes Revisited	309
10.1.2	Preemption and the Role of Structural Information	311
10.1.3	Overdetermination and Quasi-Dependence	313
10.1.4	Mackie's INUS Condition	313
10.2	Production, Dependence, and Sustainance	316
10.3	Causal Beams and Sustainance-Based Causation	318
10.3.1	Causal Beams: Definitions and Implications	318
10.3.2	Examples: From Disjunction to General Formulas	320
10.3.3	Beams, Preemption, and the Probability of Single-Event Causation	322
10.3.4	Path-Switching Causation	324
10.3.5	Temporal Preemption	325
10.4	Conclusions	327

Epilogue The Art and Science of Cause and Effect	
A public lecture delivered November 1996 as part of the UCLA Faculty Research Lectureship Program	331
<i>Bibliography</i>	359
<i>Name Index</i>	375
<i>Subject Index</i>	379