

Contents

<i>Preface</i>	<i>page xi</i>
1 Introduction	I
1.1 Causal Inference and Causal Discovery	2
1.2 Research Questions that Can Be Answered with Direction Dependence Analysis	4
1.2.1 Distinguishing Between Cause and Effect	5
1.2.2 Identifying the Presence of Hidden Confounders	6
1.2.3 Identifying Data Sectors of Predominant Causal Models	7
1.2.4 Identifying Mechanisms of Mediation	8
1.2.5 How Robust Are the Causal Claims?	8
1.3 Basics of Distributions of Variables	9
1.3.1 Types of Moments of Variables	9
1.3.2 Mean	10
1.3.3 Variance	12
1.3.4 Skewness	13
1.3.5 Kurtosis	16
1.3.6 Higher than Fourth Moments	19
1.3.7 Cumulants	21
1.3.8 Joint Moments	22
1.4 Take-Home Messages	26
2 The Linear Regression Model	28
2.1 The General Linear Model	28
2.2 Parameter Estimation	30
2.2.1 Ordinary Least Squares Estimation	31
2.2.2 Maximum Likelihood Estimation	32
2.3 Symmetry Properties of the Linear Regression Model	33
2.4 Real-World Data Example: Aggression in Adolescence	35
2.5 Asymmetry Properties of Cause and Effect	37
2.6 Take-Home Messages	40

vi	<i>Contents</i>	
3	Asymmetry Properties of Distributions of Observed Variables	42
3.1	Asymmetry of Third Moments	44
3.1.1	Likelihood Ratio Approximations	48
3.1.2	Implications and Graphical Interpretations	51
3.1.3	Statistical Inference for Third Moments	53
3.1.4	Statistical Performance of Third Moment Measures	56
3.1.5	Data Examples	59
3.2	Asymmetry of Fourth Moments	63
3.2.1	Likelihood Ratio Approximations	66
3.2.2	Implications and Graphical Interpretations	70
3.2.3	Statistical Inference for Fourth Moments	72
3.2.4	Statistical Performance of Fourth Moment Measures	74
3.2.5	Data Examples	77
3.3	Guidelines for Causal Model Selection	81
3.4	Take-Home Messages	82
4	Asymmetry Properties of Error Distributions	85
4.1	Asymmetry of Third Moments of Errors	89
4.1.1	Causal Models with Symmetric Errors	90
4.1.2	Causal Models with Asymmetric Errors	91
4.1.3	Statistical Inference for Third Moment Measures of Errors	98
4.1.4	Statistical Performance of Third Moment Measures of Errors	101
4.1.5	Data Examples	103
4.2	Asymmetry of Fourth Moments of Errors	111
4.2.1	Causal Models with Mesokurtic Errors	112
4.2.2	Causal Models with Non-Mesokurtic Errors	113
4.2.3	Statistical Inference for Fourth Moment Measures of Errors	117
4.2.4	Statistical Performance of Fourth Moment Measures of Errors	120
4.2.5	Data Examples	123
4.3	Guidelines for Causal Model Selection	131
4.4	Take-Home Messages	134
5	Independence Properties of Causes and Errors	135
5.1	Uncorrelatedness and Stochastic Independence	136
5.2	Asymmetry of Independence Structures in Competing Causal Models	137
5.3	Measures of Independence and Statistical Inference	141
5.3.1	Nonlinear Correlation Tests	141
5.3.2	Special Case: Homoscedasticity Tests	144
5.3.3	Distance Correlation	146
5.3.4	Mutual Information	148
5.3.5	Hilbert-Schmidt Independence Criterion	150
5.4	Performance of Independence-based Causal Model Selection	152
5.4.1	Accuracy for Asymmetric Distributions	152
5.4.2	Accuracy for Symmetric Non-Gaussian Distributions	156

<i>Contents</i>	<i>vii</i>
5.5 Data Examples	160
5.5.1 Artificial Examples with Known Causal Mechanisms	160
5.5.2 Real-World Data Example: Development of Numerical Cognition	165
5.6 Take-Home Messages	170
6 Direction of Dependence under Latent Confounding	172
6.1 Causal Models with Latent Confounders	173
6.2 Causal Effect Estimation under Latent Confounding	177
6.2.1 Causal Effect Estimation Using Third Moments	177
6.2.2 Causal Effect Estimation Using Fourth Moments	179
6.3 Distributional DDA Measures under Latent Confounding	181
6.3.1 DDA Measures for Observed Variables	181
6.3.2 DDA Measures for Competing Error Terms	185
6.3.3 Performance of Distributional DDA Measures under Latent Confounding	187
6.4 Independence Measures with Latent Confounders	190
6.4.1 Performance of Independence Measures under Latent Confounding	193
6.5 Real-World Data Example: Depression and Media Exposure	196
6.6 Take-Home Messages	204
7 The Integrated Framework of Direction Dependence Analysis	206
7.1 Integrating the Components of Direction Dependence Analysis	207
7.2 Covariate Selection and Adjustment	212
7.2.1 Model Definitions	214
7.2.2 Adjusting for Covariate Effects	216
7.2.3 Evaluating the Eligibility of Covariates	217
7.3 Assumptions of Direction Dependence Analysis	219
7.3.1 Non-Gaussianity of Variables	220
7.3.2 Linearity of Variable Relation	220
7.3.3 Homogeneity of the Causal Effect	221
7.3.4 No Outliers and Influential Data Observations	222
7.3.5 No Measurement Error in the Predictor	229
7.4 Real-World Data Example: Development of Numerical Cognition	229
7.5 Take-Home Messages	236
8 Stability and Sensitivity Analyses	237
8.1 Stability Analysis using Bootstrap Aggregation	238
8.1.1 Bootstrap Aggregated Measures of Direction of Dependence	238
8.1.2 Aggregating Separate Significance Tests	239
8.1.3 Stability of Causal Model Selection	240
8.1.4 Real-World Data Example: Stability of Causal Numerical Cognition Development	241

viii	<i>Contents</i>	
8.2	Sensitivity against Latent Confounding	243
8.2.1	The Algorithm	244
8.2.2	Prototypical Sensitivity Patterns	246
8.2.3	Real-World Data Example: Sensitivity of Causal Numerical Cognition Development	249
8.3	Take-Home Messages	254
9	Extensions and Applications	255
9.1	Direction of Dependence in Moderation Models	255
9.1.1	Principles of Moderation	256
9.1.2	Causally Competing Moderation Models	259
9.1.3	Conditional Direction Dependence Analysis	260
9.1.4	Real-World Data Example: Causal Mechanism of Gambling Risk Factors	265
9.2	Direction of Dependence in Mediation Models	272
9.2.1	Elements of Mediation Analysis	273
9.2.2	Causal Assumptions of Mediation Models	274
9.2.3	Testing Causal Assumptions in Mediation Models	278
9.2.4	Real-World Data Example: Teacher Classroom Management, Student Behavior, and Academic Competence	286
9.3	Direction of Dependence in Nonlinear Models	291
9.3.1	Nonlinear Models with Additive Errors	291
9.3.2	Direction of Dependence in Linearizable Models	294
9.3.3	Real-World Data Examples	296
9.4	Take-Home Messages	301
10	Statistical Software	303
10.1	Overview of Statistical Software	303
10.2	Direction Dependence Analysis in R	304
10.2.1	Observed Variable Distributions: <code>dda.vardist()</code>	304
10.2.2	Error Distributions: <code>dda.resdist()</code>	307
10.2.3	Predictor–Error Independence: <code>dda.indep()</code>	310
10.3	Conditional Direction Dependence Analysis in R	314
10.3.1	Observed Variable Distributions: <code>cd.da.vardist()</code>	314
10.3.2	Predictor–Error Independence: <code>cd.da.indep()</code>	317
11	Concluding Remarks	320
11.1	Embedding DDA into the Landscape of Causality Research	320
11.1.1	Can DDA Replace Randomized Experiments? No	321
11.1.2	Three Application Scenarios of DDA in the Process of Causal Inquiry	322
11.2	Utilizing Non-Gaussianity in Causal Inference	324
11.2.1	Instrumental Variable Approaches	324
11.2.2	Cross-lagged Panel Models	327
11.2.3	Time Series Analysis and Granger Causality	328
11.2.4	Reciprocal Causation	330

	<i>Contents</i>	ix
11.3	Categorical Data Domains	331
11.4	Importance of Data Quality	332
11.4.1	Presence of Measurement Error	332
11.4.2	Psychometric Measurement of Variables	334
11.4.3	Other Forms of Data Contamination	335
11.5	Conclusions	336
	<i>References</i>	339
	<i>Index</i>	365