

Contents

Foreword	xi
Preface.....	xiii
Acknowledgments.....	xvii
Authors.....	xix
Chapter 1 There Is More to Assessing Risk Than Statistics	1
1.1 Introduction.....	1
1.2 Predicting Economic Growth: The Normal Distribution and Its Limitations	3
1.3 Patterns and Randomness: From School League Tables to Siegfried and Roy	7
1.4 Dubious Relationships: Why You Should Be Very Wary of Correlations and Their Significance Values	10
1.5 Spurious Correlations: How You Can Always Find a Silly ‘Cause’ of Exam Success	14
1.6 The Danger of Regression: Looking Back When You Need to Look Forward	16
1.7 The Danger of Averages.....	18
1.7.1 What Type of Average?.....	19
1.7.2 When Averages Alone Will Never Be Sufficient for Decision Making.....	20
1.8 When Simpson’s Paradox Becomes More Worrisome.....	21
1.9 Uncertain Information and Incomplete Information: Do Not Assume They Are Different	23
1.10 Do Not Trust Anybody (Even Experts) to Properly Reason about Probabilities	26
1.11 Chapter Summary	29
Further Reading	29
Chapter 2 The Need for Causal, Explanatory Models in Risk Assessment	31
2.1 Introduction.....	31
2.2 Are You More Likely to Die in an Automobile Crash When the Weather Is Good Compared to Bad?.....	31
2.3 When Ideology and Causation Collide.....	35
2.4 The Limitations of Common Approaches to Risk Assessment	37
2.4.1 Measuring Armageddon and Other Risks	37
2.4.2 Risks and Opportunities.....	39
2.4.3 Risk Registers and Heat Maps	40
2.5 Thinking about Risk Using Causal Analysis	42
2.6 Applying the Causal Framework to <i>Armageddon</i>	46
2.7 Summary	49
Further Reading	49
Chapter 3 Measuring Uncertainty: The Inevitability of Subjectivity.....	51
3.1 Introduction.....	51
3.2 Experiments, Outcomes, and Events.....	52
3.2.1 Multiple Experiments.....	56
3.2.2 Joint Experiments.....	57

3.2.3	Joint Events and Marginalization.....	58
3.3	Frequentist versus Subjective View of Uncertainty	60
3.4	Summary	67
	Further Reading.....	68
4	The Basics of Probability	69
4.1	Introduction	69
4.2	Some Observations Leading to Axioms and Theorems of Probability	69
4.3	Probability Distributions.....	81
4.3.1	Probability Distributions with Infinite Outcomes.....	83
4.3.2	Joint Probability Distributions and Probability of Marginalized Events.....	85
4.3.3	Dealing with More than Two Variables	88
4.4	Independent Events and Conditional Probability.....	89
4.5	Binomial Distribution.....	96
4.6	Using Simple Probability Theory to Solve Earlier Problems and Explain Widespread Misunderstandings	101
4.6.1	The Birthday Problem.....	101
4.6.2	The Monty Hall Problem	103
4.6.3	When Incredible Events Are Really Mundane	105
4.6.4	When Mundane Events Really Are Quite Incredible	109
4.7	Summary	110
	Further Reading	111
5	Bayes' Theorem and Conditional Probability.....	113
5.1	Introduction	113
5.2	All Probabilities Are Conditional	113
5.3	Bayes' Theorem.....	116
5.4	Using Bayes' Theorem to Debunk Some Probability Fallacies	121
5.4.1	Traditional Statistical Hypothesis Testing	122
5.4.2	The Prosecutor Fallacy Revisited	124
5.4.3	The Defendant's Fallacy	124
5.4.4	Odds Form of Bayes and the Likelihood Ratio	125
5.5	Second-Order Probability	127
5.6	Summary	129
	Further Reading	129
6	From Bayes' Theorem to Bayesian Networks.....	131
6.1	Introduction	131
6.2	A Very Simple Risk Assessment Problem	132
6.3	Accounting for Multiple Causes (and Effects)	134
6.4	Using Propagation to Make Special Types of Reasoning Possible	137
6.5	The Crucial Independence Assumptions	139
6.6	Structural Properties of BNs.....	144
6.6.1	Serial Connection: Causal and Evidential Trails	144
6.6.2	Diverging Connection: Common Cause	147
6.6.3	Converging Connection: Common Effect.....	149
6.6.4	Determining Whether Any Two Nodes in a BN Are Dependent	151

6.7	Propagation in Bayesian Networks.....	153
6.8	Using BNs to Explain Apparent Paradoxes.....	156
6.8.1	Revisiting the Monty Hall Problem	156
6.8.1.1	Simple Solution.....	156
6.8.1.2	Complex Solution	157
6.8.2	Revisiting Simpson's Paradox	161
6.9	Steps in Building and Running a BN Model.....	162
6.9.1	Building a BN Model	162
6.9.2	Running a BN Model	166
6.9.3	Inconsistent Evidence.....	168
6.10	Summary	169
	Further Reading	169
	Theoretical Underpinnings.....	169
	BN Applications	169
	Nature and Theory of Causality	170
	Uncertain Evidence (Soft and Virtual)	170
Chapter 7	Defining the Structure of Bayesian Networks	171
7.1	Introduction	171
7.2	Causal Inference and Choosing the Correct Edge Direction	172
7.3	The Idioms.....	174
7.3.1	The Cause-Consequence Idiom	175
7.3.2	Measurement Idiom	177
7.3.3	Definitional/Synthesis Idiom.....	184
7.3.3.1	Case 1: Definitional Relationship between Variables.....	184
7.3.3.2	Case 2: Hierarchical Definitions	184
7.3.3.3	Case 3: Combining Different Nodes Together to Reduce Effects of Combinatorial Explosion ("Divorcing")	185
7.3.4	Induction Idiom.....	188
7.4	The Problems of Asymmetry and How to Tackle Them	190
7.4.1	Impossible Paths.....	190
7.4.2	Mutually Exclusive Paths	192
7.4.3	Distinct Causal Pathways	194
7.4.4	Taxonomic Classification	196
7.5	Multiobject Bayesian Network Models	202
7.6	The Missing Variable Fallacy	207
7.7	Conclusions	212
	Further Reading	213
Chapter 8	Building and Eliciting Node Probability Tables	215
8.1	Introduction	215
8.2	Factorial Growth in the Size of Probability Tables.....	215
8.3	Labeled Nodes and Comparative Expressions	217
8.4	Boolean Nodes and Functions.....	221
8.4.1	The Asia Model.....	222
8.4.2	The OR Function for Boolean Nodes.....	227
8.4.3	The AND Function for Boolean Nodes	234
8.4.4	M from N Operator.....	235

8.4.5	NoisyOR Function for Boolean Nodes	236
8.4.6	Weighted Averages	241
8.5	Ranked Nodes	244
8.5.1	Background	244
8.5.2	Solution: Ranked Nodes with the TNormal Distribution	246
8.5.3	Alternative Weighted Functions for Ranked Nodes	252
8.5.4	Hints and Tips When Working with Ranked Nodes and NPTs	255
8.5.4.1	Tip 1: Use the Weighted Functions as Far as Possible	255
8.5.4.2	Tip 2: Make Use of the Fact That a Ranked Node Parent Has an Underlying Numerical Scale	255
8.5.4.3	Tip 3: Do Not Forget the Importance of the Variance in the TNormal Distribution	256
8.5.4.4	Tip 4: Change the Granularity of a Ranked Scale without Having to Make Any Other Changes	259
8.5.4.5	Tip 5: Do Not Create Large, Deep, Hierarchies Consisting of Rank Nodes	260
8.6	Elicitation	260
8.6.1	Elicitation Protocols and Cognitive Biases	260
8.6.2	Scoring Rules and Validation	263
8.6.3	Sensitivity Analysis	264
8.7	Summary	265
	Further Reading	265
Chapter 9	Numeric Variables and Continuous Distribution Functions	267
9.1	Introduction	267
9.2	Some Theory on Functions and Continuous Distributions	268
9.3	Static Discretization	273
9.4	Dynamic Discretization	280
9.5	Using Dynamic Discretization	283
9.5.1	Prediction Using Dynamic Discretization	283
9.5.2	Conditioning on Discrete Evidence	287
9.5.3	Parameter Learning (Induction) Using Dynamic Discretization	289
9.5.3.1	Classical versus Bayesian Modeling	289
9.5.3.2	Bayesian Hierarchical Model Using Beta-Binomial	294
9.6	Avoiding Common Problems When Using Numeric Nodes	300
9.6.1	Unintentional Negative Values in a Node's State Range	300
9.6.2	Potential Division by Zero	301
9.6.3	Using Unbounded Distributions on a Bounded Range	301
9.6.4	Observations with Very Low Probability	302
9.7	Summary	303
	Further Reading	303
Chapter 10	Hypothesis Testing and Confidence Intervals	305
10.1	Introduction	305
10.2	Hypothesis Testing	305
10.2.1	Bayes Factors	306
10.2.2	Testing for Hypothetical Differences	308
10.2.3	Comparing Bayesian and Classical Hypothesis Testing	311

10.2.4	Model Comparison: Choosing the Best Predictive Model	
10.2.5	Accommodating Expert Judgments about Hypotheses	
10.2.6	Distribution Fitting as Hypothesis Testing	
10.2.7	Bayesian Model Comparison and Complex Causal Hypotheses	
10.3	Confidence Intervals	
10.3.1	The Fallacy of Frequentist Confidence Intervals	
10.3.2	The Bayesian Alternative to Confidence Intervals	
10.4	Summary	
	Further Reading	
Chapter 11	Modeling Operational Risk	
11.1	Introduction	
11.2	The Swiss Cheese Model for Rare Catastrophic Events	
11.3	Bow Ties and Hazards	
11.4	Fault Tree Analysis (FTA)	
11.5	Event Tree Analysis (ETA)	
11.6	Soft Systems, Causal Models, and Risk Arguments	
11.7	KUUUB Factors	
11.8	Operational Risk in Finance	
11.8.1	Modeling the Operational Loss Generation Process	
11.8.2	Scenarios and Stress Testing	
11.9	Summary	
	Further Reading	
Chapter 12	Systems Reliability Modeling	
12.1	Introduction	
12.2	Probability of Failure on Demand for Discrete Use Systems	
12.3	Time to Failure for Continuous Use Systems	
12.4	System Failure Diagnosis and Dynamic Bayesian Networks	
12.5	Dynamic Fault Trees (DFTs)	
12.6	Software Defect Prediction	
12.7	Summary	
	Further Reading	
Chapter 13	Bayes and the Law	
13.1	Introduction	
13.2	The Case for Bayesian Reasoning about Legal Evidence	
13.3	Building Legal Arguments Using Idioms	
13.3.1	The Evidence Idiom	
13.3.2	The Evidence Accuracy Idiom	
13.3.3	Idioms to Deal with the Key Notions of "Motive" and "Opportunity"	
13.3.4	Idiom for Modeling Dependency between Different Pieces of Evidence	
13.3.5	Alibi Evidence Idiom	
13.3.6	Explaining away Idiom	
13.4	Putting it All Together: Vole Example	
13.5	Using BNs to Expose Further Fallacies of Legal Reasoning	

13.5.1 The Jury Observation Fallacy	433
13.5.2 The “Crimewatch UK” Fallacy.....	435
13.6 Summary.....	438
Further Reading.....	438
Appendix A: The Basics of Counting.....	441
Appendix B: The Algebra of Node Probability Tables.....	449
Appendix C: Junction Tree Algorithm.....	455
Appendix D: Dynamic Discretization	465
Appendix E: Statistical Distributions	483
Index.....	495