

LOC Page

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
Chapter 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
Chapter 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
Chapter 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	<i>M</i> from <i>N</i> 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.....	315
10.2.5	Accommodating Expert Judgments about Hypotheses	322
10.2.6	Distribution Fitting as Hypothesis Testing.....	325
10.2.7	Bayesian Model Comparison and Complex Causal Hypotheses	326
10.3	Confidence Intervals.....	333
10.3.1	The Fallacy of Frequentist Confidence Intervals	333
10.3.2	The Bayesian Alternative to Confidence Intervals	337
10.4	Summary	340
	Further Reading	341
Chapter 11	Modeling Operational Risk.....	343
11.1	Introduction	343
11.2	The Swiss Cheese Model for Rare Catastrophic Events	344
11.3	Bow Ties and Hazards.....	347
11.4	Fault Tree Analysis (FTA).....	348
11.5	Event Tree Analysis (ETA).....	354
11.6	Soft Systems, Causal Models, and Risk Arguments	357
11.7	KUUUB Factors.....	362
11.8	Operational Risk in Finance	364
11.8.1	Modeling the Operational Loss Generation Process	364
11.8.2	Scenarios and Stress Testing	372
11.9	Summary	375
	Further Reading	376
Chapter 12	Systems Reliability Modeling.....	377
12.1	Introduction	377
12.2	Probability of Failure on Demand for Discrete Use Systems	378
12.3	Time to Failure for Continuous Use Systems.....	380
12.4	System Failure Diagnosis and Dynamic Bayesian Networks	383
12.5	Dynamic Fault Trees (DFTs).....	387
12.5.1	Modeling Other DFT Gates	390
12.6	Software Defect Prediction	395
12.7	Summary	404
	Further Reading	404
Chapter 13	Bayes and the Law	407
13.1	Introduction	407
13.2	The Case for Bayesian Reasoning about Legal Evidence	408
13.3	Building Legal Arguments Using Idioms	411
13.3.1	The Evidence Idiom	411
13.3.2	The Evidence Accuracy Idiom.....	414
13.3.3	Idioms to Deal with the Key Notions of “Motive” and “Opportunity”	417
13.3.4	Idiom for Modeling Dependency between Different Pieces of Evidence... 420	
13.3.5	Alibi Evidence Idiom	422
13.3.6	Explaining away Idiom	425
13.4	Putting it All Together: Vole Example	428
13.5	Using BNs to Expose Further Fallacies of Legal Reasoning	433

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

Foreword

Probabilistic models based on directed acyclic graphs have a long and rich tradition, beginning with work by the geneticist Sewall Wright in the 1920s. Variants have appeared in many fields. Within statistics, such models are known as directed graphical models; within cognitive science and artificial intelligence, such models are known as Bayesian networks. The name honors the Reverend Thomas Bayes (1702–1761), whose rule for updating probabilities in the light of new evidence is the foundation of the approach. The initial development of Bayesian networks in the late 1970s was motivated by the need to model the top-down (semantic) and bottom-up (perceptual) combination of evidence in reading. The capability for bidirectional inferences, combined with a rigorous probabilistic foundation, led to the rapid emergence of Bayesian networks as the method of choice for uncertain reasoning in AI and expert systems, replacing earlier ad-hoc rule-based schemes. Perhaps the most important aspect of Bayesian networks is that they are direct representations of the world, not of reasoning processes. The arrows in the diagrams represent real causal connections and not the flow of information during reasoning (as in rule-based systems or neural networks). Reasoning processes can operate on Bayesian networks by propagating information in any direction. For example, if the sprinkler is on, then the pavement is probably wet (prediction); if someone slips on the pavement, that also provides evidence that it is wet (abduction, or reasoning to a probable cause). On the other hand, if we see that the pavement is wet, that makes it more likely that the sprinkler is on or that it is raining (abduction); but if we then observe that the sprinkler is on, that reduces the likelihood that it is raining. It is the ability to perform this last form of reasoning—called explaining away—that makes Bayesian networks so powerful compared to rule-based systems or neural networks. They are especially useful and important for risk assessment and decision-making.

Although Bayesian networks are now used widely in many disciplines, those responsible for developing (as opposed to using) Bayesian network models typically require highly specialist knowledge of mathematics, probability, statistics, and computing. Part of the reason for this is that, although there have been several excellent books dedicated to Bayesian Networks and related methods, these books tend to be aimed at readers who already have a high level of mathematical sophistication—typically they are books that would be used at graduate or advanced undergraduate level in mathematics, statistics, or computer science. As such they are not accessible to readers who are not already proficient in those subjects. This book is an exciting development because it addresses this problem. While I am sure it would be suitable for undergraduate courses on probability and risk, it should be understandable by any numerate reader interested in risk assessment and decision making. The book provides sufficient motivation and examples (as well as the mathematics and probability where needed from scratch) to enable readers to understand the core principles and power of Bayesian networks. However, the focus is on ensuring that readers can build practical Bayesian network models, rather than understand in depth the underlying algorithms and theory. Indeed readers are provided with a tool that performs the propagation, so they will be able to build their own models to solve real-world risk assessment problems.

Judea Pearl

UCLA Computer Science Department

Los Angeles, California

Preface

Businesses and governments must often assess and manage risk in areas where there is little or no direct historical data to draw upon, or where relevant data is difficult to identify. The international financial crisis that started in 2008 was not predicted by the majority of the world's leading economists, because they relied on models based on historical statistical data that could not adapt to new circumstances even when those circumstances (in this case the collapse of the mortgage subprime market) were foreseeable by some experts. In short, analysts are too dependent on models that are good for predicting the past but poor at predicting the future.

The challenges are similarly acute when the source of the risk is novel: terrorist attacks, ecological disasters, major project failures, and more general failures of novel systems, marketplaces, and business models. Even though we may have little or no historical data in these cases, there is often an abundance of expert (but subjective) judgment, as well as diverse information and data on indirectly related risks. These are the types of situations that can be successfully addressed using Bayesian networks (BNs), even when classical, data-driven approaches to risk assessment are not possible.

BNs describe networks of causes and effects, using a graphical framework that provides for the rigorous quantification of risks and the clear communication of results. They can combine historical data and expert judgment, using calculations that are based on a theorem created by the Reverend Thomas Bayes dating back to 1763. The theorem provides the only rational and consistent way to solve the problem of how to update a belief in some uncertain event (such as a decline in a company's share price of more than 10% next year) when we observe new evidence about the event (such as better than expected first quarter figures). The problem of correctly updating beliefs in the light of new evidence is one that is central to many disciplines (law, medicine, and engineering as well as finance). It is also prevalent in our everyday decision making (even those of us who are not gamblers).

BNs are now widely recognized as an exciting and powerful technology for handling risk assessment, uncertainty, and decision making. During the last decade, researchers have incorporated BN techniques into easy-to-use toolsets, which in turn have enabled the development of decision support systems in a diverse set of application domains, including medical diagnosis, safety assessment, the law, forensics, procurement, equipment fault diagnosis and software quality. Further technology and tool advancements mean that end users, rather than just researchers, are now able to develop and deploy their own BN-based solutions. Recent commercial case studies provide evidence of impressive returns on investment from these techniques.

But, although many BN applications are now in everyday use, BNs have not yet achieved the mainstream penetration that the approach deserves. Although there have been several excellent books dedicated to BNs and related methods (these are described in the Further Reading sections), these books are largely aimed at the mathematical and statistical community and focus more on algorithms and theory rather than practical real-world problem solving and model building. Hence, there is a shortage of material that is truly accessible to people who would benefit from BN solutions but who are not already qualified mathematicians or statisticians.

This book aims to fix this problem. Although it is suitable for undergraduate courses on probability and risk, it is written to be understandable by other professional people generally interested in risk assessment and decision making. Unlike other BN books, this one makes no assumptions about previous probability and statistical knowledge. It is driven by real examples that are introduced early on as motivation, with the probability and statistics introduced (from scratch) as and when necessary. The more mathematical topics are separated from the main text by comprehensive use of boxes and appendices. The focus is on applications and practical model building. A free version of the powerful commercial software tool AgenaRisk is available for those who have purchased the book, as we think the only real way to learn about BNs is to build and use them.

Many of the examples in this book are influenced by our academic research but also by our experience in putting the ideas into practice with commercial and government decision and policy makers. Together we have consulted for a wide variety of commercial and government organizations, including Absa Bank, Bosch, Ericsson, Motorola, NATS, Philips, QinetiQ, Dstl (UK Ministry of Defence), Royal Bank of Canada, TNO, and VocaLink. We are firm believers in technology transfer and putting useful decision-making theory into the hands of those at the sharp end.

The book has two parts. The first ten chapters of the book teach all the basics of probability and risk, and about building and using BN models, whereas the last three chapters go into the detailed applications. The underlying BN algorithms appear in appendices rather than the main text since there is no need to understand these to build and use BN models. Although purists have argued that only by understanding the algorithms can you understand the limitations and hence build efficient BN models, we overcome this by providing pragmatic advice about model building to ensure models are built efficiently. Our approach means that the main body of the text is free of the intimidating mathematics that has been a major impediment to the more widespread use of BNs.

The book has a dedicated Website, www.BayesianRisk.com, which contains executable versions of all of the models described, exercises and worked solutions for all chapters, PowerPoint slides, numerous other resources, and a free downloadable copy of the AgenaRisk software to readers who have purchased a copy of the book.

For our parents, wives and children.

Acknowledgments

Numerous people have helped in various ways to ensure that we achieved the best possible outcome for this book. We would especially like to thank the following:

Colin Aitken, Carol Alexander, Peter Ayton, David Balding, Beth Bateman, George Bearfield, Daniel Berger, Nic Birtles, Robin Bloomfield, Bill Boyce, Rob Calver, Neil Cattle, Patrick Cates, Chris Chapman, Xiaoli Chen, Keith Clarke, Julie Cooper, Robert Cowell, Anthony Constantinou, Paul Curzon, Phil Dawid, Chris Eagles, Shane Cooper, Eugene Dementiev, Itiel Dror, John Elliott, Phil Evans, Ian Evett, Geir Fagerhus, Simon Forey, Duncan Gillies, Jean-Jacques Gras, Gerry Graves, David Hager, George Hanna, David Hand, Roger Harris, Peter Hearty, Joan Hunter, Jose Galan, Steve Gilmour, Shlomo Gluck, James Gralton, Richard Jenkinson, Adrian Joseph, Ian Jupp, Agnes Kaposi, Paul Kaye, Kevin Korb, Paul Krause, Dave Lagnado, Helge Langseth, Steffen Lauritzen, Robert Leese, Peng Lin, Bev Littlewood, Paul Loveless, Peter Lucas, Bob Malcom, Amber Marks, David Marquez, William Marsh, Peter McOwan, Tim Menzies, Phil Mercy, Martin Newby, Richard Nobles, Magda Osman, Max Parmar, Judea Pearl, Elena Perez-Minana, Andrej Peitschker, Ursula Martin, Shoaib Qureshi, Lukasz Radlinksy, Soren Riis, Edmund Robinson, Thomas Roelleke, Angela Saini, Thomas Schulz, Jamie Sherrah, Leila Schneps, David Schiff, Bernard Silverman, Adrian Smith, Ian Smith, Jim Smith, Julia Sonander, David Spiegelhalter, Andrew Stuart, Alistair Sutcliffe, Lorenzo Strigini, Nigel Tai, Manesh Tailor, Franco Taroni, Ed Tranham, Marc Trepanier, Keith van Rijsbergen, Richard Tonkin, Sue White, Robin Whitty, Rosie Wild, Patricia Wiltshire, Rob Wirszyz, David Wright and Barbaros Yet.