# International Series in Operations Research & Management Science

Volume 270



More information about this series at http://www.springer.com/series/6161

Louis Anthony Cox Jr. • Douglas A. Popken
Richard X. Sun

# Causal Analytics for Applied Risk Analysis

Springer

Louis Anthony Cox Jr.
Cox Associates
Denver, CO, USA

Douglas A. Popken
Cox Associates
Littleton, CO, USA

Richard X. Sun
Cox Associates
East Brunswick, NJ, USA

ISSN 0884-8289 ISSN 2214-7934 (electronic)
International Series in Operations Research & Management Science
ISBN 978-3-319-78240-9 ISBN 978-3-319-78242-3 (eBook)
https://doi.org/10.1007/978-3-319-78242-3

Library of Congress Control Number: 2018937974



Printed on acid-free paper

This Springer imprint is published by the registered company Springer International Publishing AG part of Springer Nature.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

*To*
*Christine and Emeline*

# Preface

Individual, group, organizational, and public policy decisions are often disconcertingly ineffective. They produce unintended and unwanted consequences, or fail to produce intended ones, even after large expenditures of hope, time, and resources. The difficulty of achieving unambiguous successes, in which costly actions or policies produce large, clear net benefits that even those who initially doubted find compelling after the fact, has been noted in areas as varied as personal financial decisions, corporate business decisions, engineering infrastructure decisions, non-profit initiatives for poverty disruption or delinquency prevention, public health efforts to curb the emergence of antibiotic-resistant superbugs, and regulation of pollutants to improve public and occupational health.

This book is about how to make more effective decisions—that is, decisions that are more likely to cause preferred outcomes and to avoid undesirable ones—by understanding and fixing what so often goes wrong. We believe that the most common reason for disappointing results from well-intended policies and actions is inadequate understanding of the causal relationships between actions and probabilities of outcomes. Actions guided by traditional statistical analyses of association patterns in observational data, such as regression modeling or epidemiological estimates of relative risk ratios, usually cannot be relied on to achieve their objectives because these traditional methods of analysis are usually not adequate for determining how changing some variables will change others. But that is what decision makers must know to make well-informed choices about what changes to implement. This book is therefore devoted to causal analytics methods that can provide answers to the crucial causal question of how changing decision variables—the things that a decision maker or policy maker can control or choose—changes probabilities of various outcomes. It presents and illustrates models, algorithms, principles, and software for deriving causal models from data and for using them to optimize decisions, evaluate effects of policies or interventions, make probabilistic predictions of the values of as-yet unobserved quantities from available data, and identify the most likely explanations for observed outcomes, including surprises and anomalies.

The first two chapters survey modern analytics methods, focusing mainly on techniques useful for decision, risk, and policy analysis. They emphasize how causal models are used throughout the rest of risk analytics in detecting and describing meaningful and useful patterns in data; predicting outcome probabilities if different courses of action are followed; identifying and prescribing a best course of action for making preferred outcomes more probable; evaluating the effects of current or past policies and interventions; and learning from experience, either individually or collaboratively, how to make choices that increase the probabilities of preferred outcomes. Chapter 2 also introduces the Causal Analytics Toolkit (CAT), a free in-browser set of analytics software tools available at http://cox-associates.com/CloudCAT, to allow readers to perform the analyses described or to apply modern analytics methods to their own data sets. Chapters 3 through 11 illustrate the application of causal analytics and risk analytics to practical risk analysis challenges, mainly related to public and occupational health risks from pathogens in food or from pollutants in air. Chapters 12 through 15 turn to broader questions of how to improve risk management decision-making by individuals, groups, organizations, institutions, and multi-generation societies with different cultures and norms for cooperation. They examine organizational learning, social risk management, and intergenerational collaboration and justice in managing risks and hazards.

Throughout the book, our main focus is on introducing and illustrating practical methods of causal modeling and analytics that practitioners can apply to improve understanding of how choices affect probabilities of consequences and, based on this understanding, to recommend choices that are more likely to accomplish their intended objectives. We believe that the analytics and big data revolutions now underway will become much more valuable as methods and software for causal analytics become more widely used to better understand how actions and policies affect outcomes.

Denver, CO, USA — Louis Anthony Cox Jr.
Littleton, CO, USA — Douglas A. Popken
East Brunswick, NJ, USA — Richard X. Sun

# Acknowledgments

This book has grown out of efforts over the past decade to understand and explain how to use data and algorithms to determine as accurately, objectively, and reproducibly as possible the effects caused by changes in decisions, actions, or policies. This quest has been inspired, encouraged, and supported by many people and organizations. It is a pleasure to thank them.

Chapters 1 and 2 are based largely on a pedagogical approach developed to teach nonspecialists about information-based causal analytics methods quickly, as part of professional development and academic graduate courses taught by Tony Cox in 2017. These included a course on Decision Analysis at the University of Colorado at Denver and professional courses at the annual meetings of the Society for Benefit Cost Analysis (SBCA), the Society for Epidemiological Research (SER), and the American Industrial Hygiene Association (AIHce). Bruce Copley, Dennis Devlin, Dale Drysdale, Susan Dudley, Gary Kochenberger, and Deborah Kellog believed in and enthusiastically supported development of a teaching approach and course materials that sought to make key ideas and methods of modern causal analytics accessible to a wider audience. We thank them.

The approach taken in those courses and in Chap. 2 of this book emphasizes the concepts and principles behind current causal analytics algorithms using a minimum of specialist jargon and mathematical notation, and then makes algorithms themselves readily available through software that can be used without learning the underlying R or Python languages and packages. Course materials are available at these links:

- https://www.aiha.org/events/AIHce2017/Documents/PDC%20Handouts%202017/PDC%20604%20Handout.pdf
- http://cox-associates.com/CausalAnalytics/

The Causal Analytics Toolkit (CAT) software and an explanation of its goals are available at these links:

- http://cox-associates.com/CloudCAT

- https://regulatorystudies.columbian.gwu.edu/causal-analytics-toolkit-cat-assessing-potential-causal-relations-data

Initial funding for development of CAT was provided by the George Washington University Regulatory Studies Center. Subsequent development of its Predictive Analytics Toolkit (PAT) module, discussed in Chap. 2, and a port from an Excel add-in version to a cloud-based version were supported in part by the American Chemistry Council. We thank Susan Dudley of the GWU Regulatory Studies Center and Rick Becker of the American Chemistry Council for their support and vision in making free, high-quality analytics software available to interested users via CAT.

The applications, ideas, and principles in Chaps. 3–15 are based mainly on recent journal articles. Material from the following articles has been used with the kind permission of Wiley-Blackwell, the publishers of *Risk Analysis: An International Journal.*

- Cox LA Jr, Popken DA. Quantitative assessment of human MRSA risks from swine. *Risk Analysis*. 2014 Sep;34(9):1639-50 (Chap. 6)
- Cox LA Jr. Overcoming learning-aversion in evaluating and managing uncertain risks. *Risk Analysis*. 2015 Oct; 35(10) (Chap. 12). (Thanks to Jim Hammitt and Lisa Robinson for a fascinating workshop at the Harvard Center for Risk Analysis that stimulated this work.)
- Paté-Cornell E, Cox LA Jr. Improving risk management: from lame excuses to principled practice. *Risk Analysis*. 2014 Jul;34(7):1228–39. (Chap. 13)

Material from the following articles has been used with the kind permission of their publishers:

- Popken DA, Cox LA Jr. Quantifying human health risks caused by Toxoplasmosis from open system production of swine. *Human and Ecological Risk Assessment*. 2015 Oct 3; 21(7): 1717–1735. (Chap. 7)
- Cox LA Jr., Popken DA, Kaplan AM, Plunkett LM, Becker RA. How well can in vitro data predict in vivo effects of chemicals? Rodent carcinogenicity as a case study. Regulatory Toxicology and Pharmacology. 2016 Jun;77:54–64. (Chap. 8)
- Cox, LA Jr, Popken DA. Has reducing PM2.5 and ozone caused reduced mortality rates in the United States? *Annals of Epidemiology*. 2015 Mar; 25 (3):162–73. (Chap. 10)
- Cox LA Jr. How accurately and consistently do laboratories measure workplace concentrations of respirable crystalline silica? *Regul Toxicol Pharmacol.* 2016 Nov;81:268–274. (Chap. 11)
- Cox T. Uncertain causation, regulation, and the courts. *Supreme Court Economic Review*. (In press.) (Chap. 14)
- Cox LA Jr., Cox ED. (2016) Intergenerational Justice in Protective and Resilience Investments with Uncertain Future Preferences and Resources. Chapter 12 in P. Gardoni, C. Murphy, and A. Rowell (Eds). *Risk Analysis of Natural Hazards: Interdisciplinary Challenges and Integrated Solutions*. Springer. New York, New York. (Chap. 15)

We thank the publishers and coauthors of these works.

Discussions with Ron Josephson of the United States Environmental Protection Agency (EPA) in the context of reviewing research proposals on health effects of air pollution helped to inspire the idea of applying causal analysis methods to determine value of information in causal networks (Chap. 2). We thank Dennis Devlin and Bruce Copley of Exxon-Mobil and Will Ollison of the American Petroleum Institute for stimulating conversations and their unswerving commitment to discovering objective scientific truth from data to inform causally effective decision and policies. As we have developed and applied software to help automatically discover scientific truth about causality from data, we have found that this approach to pursuing more objective and reliable scientific inference is not always welcome. Advocates of expert judgment-based and modeling assumption-based approaches to causal inference in risk assessment have not always embraced the ideas that computer algorithms can now be far more accurate and objective than human experts in discovering true causal relations in data, and in identifying and rejecting false causal hypotheses; and that modeling judgments and expert interpretations of statistical patterns are not necessary or desirable for drawing valid causal inferences from data. We expect this analytics-centric perspective to continue to grow in popularity as causal discovery algorithms prove their value in a wide array of risk analysis applications. Meanwhile, we thank the visionaries who are pushing to make automated, objective, reproducible, algorithmic approaches to causal model discovery and validation a practical reality.

Finally, we thank Douglas Hubbard of Hubbard Decision Research for inviting lectures and discussions of the causal analytics framework in Chap. 2 at the American Statistical Association Symposium on Statistical Inference (*Scientific Method for the 21st Century: A World Beyond* $p < 0.05$ in October of 2017) and Seth Guikema of the University of Michigan for inviting the 2017 Wilbert Steffy Distinguished Lecture on Causal Analytics for Risk Management: Making Advanced Analytics More Useful at the University of Michigan Department of Industrial Engineering and Operations Research in November of 2017. The opportunity to prepare and present these lectures and to participate in the very stimulating discussions that followed contributed to the final exposition in Chaps. 1 and 2.

# Contents