Elements Of Causal Inference Foundations And Learning Algorithms Adaptive Computation And Machine Learning Series

The application of causal inference methods is growing exponentially in fields that deal with observational data. Written by pioneers in the field, this practical book presents an authoritative yet accessible overview of the methods and applications of causal inference. With a wide range of detailed, worked examples using real epidemiologic data as well as software for replicating the analyses, the text provides a thorough introduction to the basics of the theory for non-time-varying treatments and the generalization to complex longitudinal data.

The classic work on qualitative methods in political science Designing Social Inquiry presents a unified approach to qualitative and quantitative research in political science, showing how the same logic of inference underlies both. This stimulating book discusses issues related to framing research questions, measuring the accuracy of data and the uncertainty of empirical inferences, discovering causal effects, and getting the most out of qualitative research. It addresses topics such as interpretation and inference, comparative case studies, constructing causal theories, dependent and explanatory variables, the limits of random selection, selection bias, and errors in measurement. The book only uses mathematical notation to clarify concepts, and assumes no prior knowledge of mathematics or statistics. Featuring a new preface by Robert O. Keohane and Gary King, this edition makes an influential work available to new generations of qualitative researchers in the social sciences.

A Turing Award-winning computer scientist and statistician shows how understanding causality has revolutionized science and will revolutionize artificial intelligence "Correlation is not causation." This mantra, chanted by scientists for more than a century, has led to a virtual prohibition on causal talk. Today, that taboo is dead. The causal revolution, instigated by Judea Pearl and his colleagues, has cut through a century of confusion and established causality -- the study of cause and effect -- on a firm scientific basis. His work explains how we can know easy things, like whether it was rain or a sprinkler that made a sidewalk wet; and how to answer hard questions, like whether a drug cured an illness. Pearl's work enables us to know not just whether one thing causes another: it lets us explore the world that is and the worlds that could have been. It shows us the essence of human thought and key to artificial intelligence. Anyone who wants to understand either needs The Book of Why.

Many racial and ethnic groups in the United States, including blacks, Hispanics, Asians, American Indians, and others, have historically faced severe discriminationâ€"pervasive and open denial of civil, social, political, educational, and economic opportunities. Today, large differences among racial and ethnic groups continue to exist in employment, income and wealth, housing, education, criminal justice, health, and other areas. While many factors may contribute to such differences, their size and extent suggest that various forms of discriminatory treatment persist in U.S. society and serve to undercut the achievement of equal opportunity. Measuring Racial Discrimination considers the definition of race and racial discrimination, reviews the existing techniques used to measure racial discrimination, and identifies new tools and areas for future research. The book conducts a thorough evaluation of current methodologies for a wide range of circumstances in which racial discrimination may occur, and makes recommendations on how to better assess the presence and effects of discrimination.

A one-of-a-kind guide to identifying and dealing with modern statistical developments in causality Written by a group of well-known experts, Statistics and Causality: Methods for Applied Empirical Research focuses on the most up-to-date developments in statistical methods in respect to causality. Illustrating the properties of statistical methods to theories of causality, the book features a summary of the latest developments in methods for statistical analysis of causality hypotheses. The book is divided into five accessible and independent parts. The first part introduces the foundations of causal structures and discusses issues associated with standard mechanistic and difference-making theories of causality. The second part features novel generalizations of methods designed to make statements concerning the direction of effects. The third part illustrates advances in Granger-causality testing and related issues. The fourth part focuses on counterfactual approaches and propensity score analysis. Finally, the fifth part presents designs for causal inference with an overview of the research designs commonly used in epidemiology. Statistics and Causality: Methods for Applied Empirical Research also includes: New statistical methodologies and approaches to causal analysis in the context of the continuing development of philosophical theories End-of-chapter bibliographies that provide references for further discussions and additional research topics Discussions on the use and applicability of software when appropriate Statistics and Causality: Methods for Applied Empirical Research is an ideal reference for practicing statisticians, applied mathematicians, psychologists, sociologists, logicians, medical professionals, epidemiologists, and educators who want to learn more about new methodologies in causal analysis. The book is also an excellent textbook for graduate-level courses in causality and qualitative logic.

A general framework for constructing and using probabilistic models of complex systems that would enable a computer to use available information for making decisions. Most tasks require a person or an automated system to reason—to reach conclusions based on available information. The framework of probabilistic graphical models, presented in this book, provides a general approach for this task. The approach is model-based, allowing interpretable models to be constructed and then manipulated by reasoning algorithms. These models can also be learned automatically from data, allowing the approach to be used in cases where manually constructing a model is difficult or even impossible. Because uncertainty is an inescapable aspect of most real-world applications, the book focuses on probabilistic models, which make the uncertainty explicit and provide models that are more faithful to reality. Probabilistic Graphical Models discusses a variety of models, spanning Bayesian networks, undirected Markov networks, discrete and continuous models, and extensions to deal with dynamical systems and relational data. For each class of models, the text describes the three fundamental cornerstones: representation, inference, and learning, presenting both basic concepts and advanced techniques. Finally, the book considers the use of the proposed framework for causal reasoning and decision making under uncertainty. The main text in each chapter provides the detailed technical development of the key ideas. Most chapters also include boxes with additional material: skill boxes, which describe techniques; case study boxes, which discuss empirical cases related to the approach described in the text, including applications in computer vision, robotics, natural language understanding, and computational biology; and concept boxes, which present significant concepts drawn from the material in the chapter. Instructors (and readers) can group chapters in various combinations, from core topics to more technically advanced material, to suit their particular needs.

Many of the concepts and terminology surrounding modern causal inference can be quite intimidating to the novice. Judea Pearl presents a book ideal for beginners in statistics, providing a comprehensive introduction to the field of causality. Examples from classical statistics are presented throughout to demonstrate the need for causality in resolving decision-making dilemmas posed by data. Causal methods are also compared to traditional statistical methods, whilst questions are provided at the end of each section to aid student learning.

Unlock today's statistical controversies and irreproducible results by viewing statistics as probing and controlling errors.

A Primer on Molecular Biology. A Primer on Kernel Methods. Support Vector Machine Applications in Computational Biology. Inexact Matching String Kernels for Protein Classification. Fast Kernels for String and Tree Matching. Local Alignment Kernels for Biological Sequences. Kernels for Graphs. Diffusion Kernels. A Kernel for Protein Secondary Structure Prediction. Heterogeneous Data Comparsion and Gene Selection with Kernel Canonical Correlation Analysis. Kernel-Based Integration of Genomic Data Using Semdefinite Programming. Protein Classification via Kernel Matrix Completion. Accurate Splice Site Detection for Caenorhabditid elegans. Gene Espression Analysis: Joint Feature Selection and Classifier Design. Gene Selection for Microarray Data.

Games that show how mathematics can solve the apparently unsolvable. This book presents a series of engaging games that seem unsolvable--but can be solved when they are translated into mathematical terms. How can players find their ID cards when the cards are distributed randomly among twenty boxes? By applying the theory of permutations. How can a player guess the color of her own hat when she can only see other players' hats? Hamming codes, which are used in communication technologies. Like magic, mathematics solves the apparently unsolvable. The games allow readers, including university students or anyone with high school-level math, to experience the joy of mathematical discovery.

This book presents a method for bringing data analysis and statistical technique into line with theory. The author begins by describing the elaboration model for analyzing the empirical association between variables. She then introduces a new concept into this model, the focal relationship. Building upon the focal relationship as the cornerstone for all subsequent analysis, two analytic strategies are developed to establish its internal validity: an exclusionary strategy to eliminate alternative explanations, and an inclusive strategy which looks at the interconnected set of relationships predicted by theory. Using real examples of social research, the author demonstrates the use of this approach for two common forms of analysis, multiple linear regression and logistic regression. Whether learning data analysis for the first time or adding new techniques to your repertoire, this book provides an excellent basis for theory-based data analysis.

Theory, algorithms, and applications of machine learning techniques to overcome “covariate shift” non-stationarity. As the power of computing has grown over the past few decades, the field of machine learning has advanced rapidly in both theory and practice. Machine learning methods are usually based on the assumption that the data generation mechanism does not change over time. Yet real-world applications of machine learning, including image recognition, natural language processing, speech recognition, robot control, and bioinformatics, often violate this common assumption. Dealing with non-stationarity is one of modern machine learning's greatest challenges. This book focuses on a specific non-stationary environment known as covariate shift, in which the distributions of inputs (queries) change but the conditional distribution of outputs (answers) is unchanged, and presents machine learning theory, algorithms, and applications to overcome this variety of non-stationarity. After reviewing the state-of-the-art research in the field, the authors discuss topics that include learning under covariate shift, model selection, importance estimation, and active learning. They describe such real world applications of covariate shift adaption as brain-computer interface, speaker identification, and age prediction from facial images. With this book, they aim to encourage future research in machine learning, statistics, and engineering that strives to create truly autonomous learning machines able to learn under non-stationarity.

Outside of randomized experiments, association does not imply causation, and yet there is nothing defective about our knowledge that smoking causes lung cancer, a conclusion reached in the absence of randomized experimentation with humans. How is that possible? If observed associations do not identify causal effects in observational studies, how can a sequence of such associations become decisive? Two or more associations may each be susceptible to unmeasured biases, yet not susceptible to the same biases. An observational study has two evidence factors if it provides two comparisons susceptible to different biases that may be combined as if from independent studies of different data by different investigators, despite using the same data twice. If the two factors concur, then they may exhibit greater insensitivity to unmeasured biases than either factor exhibits on its own. Replication and Evidence Factors in Observational Studies includes four parts: A concise introduction to causal inference, making the book self-contained Practical examples of evidence factors from the health and social sciences with analyses in R The theory of evidence factors Study design with evidence factors A companion R package evident is available from CRAN.

Ryall and Bramson's Inference and Intervention is the first textbook on causal modeling with Bayesian networks for business applications. In a world of resource scarcity, a decision about which business elements to control or change – as the authors put it, a managerial intervention – must precede any decision on how to control or change them, and understanding causality is crucial to making effective interventions. The authors cover the full spectrum of causal modeling techniques useful for the managerial role, whether for intervention, situational assessment, strategic decision-making, or forecasting. From the basic concepts and nomenclature of causal modeling to decision tree analysis, qualitative methods, and quantitative modeling tools, this book offers a toolbox for MBA students and business professionals to make successful decisions in a managerial setting.

In the sixteenth and seventeenth centuries, gamblers and mathematicians transformed the idea of chance from a mystery into the discipline of probability, setting the stage for a series of breakthroughs that enabled or transformed innumerable fields, from gambling, mathematics, statistics, economics, and finance to physics and computer science. This book tells the story of ten great ideas about chance and the thinkers who developed them, tracing the philosophical implications of these ideas as well as their mathematical impact.

A fundamental book for social researchers. It provides a first-class, reliable guide to the basic issues in data analysis. Scholars and students can turn to it for teaching and applied needs with confidence.

Written by one of the preeminent researchers in the field, this book provides a comprehensive exposition of modern analysis of causation. It shows how causality has grown from a nebulous concept into a mathematical theory with significant applications in the fields of statistics, artificial intelligence, economics, philosophy, cognitive science, and the health and social sciences. Judea Pearl presents and unifies the probabilistic, manipulative, counterfactual, and structural approaches to causation and devises simple mathematical tools for studying the relationships between causal connections and statistical associations. Cited in more than 2,100 scientific publications, it continues to liberate scientists from the traditional molds of statistical thinking. In this revised edition, Judea Pearl elucidates thorny issues, answers readers' questions, and offers a panoramic view of recent advances in this field of research. Causality will be of interest to students and professionals in a wide variety of fields. Dr Judea Pearl has received the 2011 Rumelhart Prize for his leading research in Artificial Intelligence (AI) and systems from The Cognitive Science Society.

One of the primary motivations for clinical trials and observational studies of humans is to infer cause and effect. Disentangling causation from confounding is of utmost importance. Fundamentals of Causal Inference explains and relates different methods of confounding adjustment in terms of potential outcomes and graphical models, including standardization, difference-in-differences estimation, the front-door method, instrumental variables estimation, and propensity score methods. It also covers effect-measure modification, precision variables, mediation analyses, and time-dependent confounding. Several real data examples, simulation studies, and analyses using R motivate the methods throughout. The book assumes familiarity with basic statistics and probability, regression, and R and is suitable for seniors or graduate students in statistics, biostatistics, and data science as well as PhD students in a wide variety of other disciplines, including epidemiology, pharmacy, the health sciences, education, and the social, economic, and behavioral sciences. Beginning with a brief history and a review of essential elements of probability and statistics, a unique feature of the book is its focus on real and simulated datasets with all binary variables to reduce complex methods down to their fundamentals. Calculus is not required, but a willingness to tackle mathematical notation, difficult concepts, and intricate logical arguments is essential. While many real data examples are included, the book also features the Double What-If Study, based on simulated data with known causal mechanisms, in the belief that the methods are best understood in circumstances where they are known to either succeed or fail. Datasets, R code, and solutions to odd-numbered exercises are available at www.routledge.com.

Did mandatory busing programs in the 1970s increase the school achievement of disadvantaged minority youth? Does obtaining a college degree increase an individual's labor market earnings? Did the use of the butterfly ballot in some Florida counties in the 2000 presidential election cost Al Gore votes? If so, was the number of miscast votes sufficiently large to have altered the election outcome? At their core, these types of questions are simple cause-and-effect questions. Simple cause-and-effect questions are the motivation for much empirical work in the social sciences. This book presents a model and set of methods for causal effect estimation that social scientists can use to address causal questions such as these. The essential features of the counterfactual model of causality for observational data analysis are presented with examples from sociology, political science, and economics.

Presents the Terminology and Methods of Mendelian Randomization for Epidemiological StudiesMendelian randomization uses genetic instrumental variables to make inferences about causal effects based on observational data. It, therefore, can be a reliable way of assessing the causal nature of risk factors, such as biomarkers, for a wide range of disea

This open access book constitutes the proceedings of the 22nd International Conference on Foundations of Software Science and Computational Structures, FOSSACS 2019, which took place in Prague, Czech Republic, in April 2019, held as part of the European Joint Conference on Theory and Practice of Software, ETAPS 2019.The 29 papers presented in this volume were carefully reviewed and selected from 85 submissions. They deal with foundational research with a clear significance for software science.

This book is intended for anyone, regardless of discipline, who is interested in the use of statistical methods to help obtain scientific explanations or to predict the outcomes of actions, experiments or policies. Much of G. Udny Yule's work illustrates a vision of statistics whose goal is to investigate when and how causal influences may be reliably inferred, and their comparative strengths estimated, from statistical samples. Yule's enterprise has been largely replaced by Ronald Fisher's conception, in which there is a fundamental cleavage between experimental and non experimental inquiry, and statistics is largely unable to aid in causal inference without randomized experimental trials. Every now and then members of the statistical community express misgivings about this turn of events, and, in our view, rightly so. Our work represents a return to something like Yule's conception of the enterprise of theoretical statistics and its potential practical benefits. If intellectual history in the 20th century had gone otherwise, there might have been a discipline to which our work belongs. As it happens, there is not. We develop material that belongs to statistics, to computer science, and to philosophy; the combination may not be entirely satisfactory for specialists in any of these subjects. We hope it is nonetheless satisfactory for its purpose.

This book summarizes recent advances in causal inference and underscores the paradigmatic shifts that must be undertaken in moving from traditional statistical analysis to causal analysis of multivariate data. Special emphasis is placed on the assumptions that underlie all causal inferences, the languages used in formulating those assumptions, the conditional nature of all causal and counterfactual claims, and the methods that have been developed for the assessment of such claims. These advances are illustrated using a general theory of causation based on the Structural Causal Model (SCM), which subsumes and unifies other approaches to causation, and provides a coherent mathematical foundation for the analysis of causes and counterfactuals. In particular, the paper surveys the development of mathematical tools for inferring (from a combination of data and assumptions) answers to three types of causal queries: those about (1) the effects of potential interventions, (2) probabilities of counterfactuals, and (3) direct and indirect effects (also known as "mediation"). Finally, the paper defines the formal and conceptual relationships between the structural and potential-outcome frameworks and presents tools for a symbiotic analysis that uses the strong features of both. The tools are demonstrated in the analyses of mediation, causes of effects, and probabilities of causation.

A hands-on approach to tasks and techniques in data stream mining and real-time analytics, with examples in MOA, a popular freely available open-source software framework. Today many information sources—including sensor networks, financial markets, social networks, and healthcare monitoring—are so-called data streams, arriving sequentially and at high speed. Analysis must take place in real time, with partial data and without the capacity to store the entire data set. This book presents algorithms and techniques used in data stream mining and real-time analytics. Taking a hands-on approach, the book demonstrates the techniques using MOA (Massive Online Analysis), a popular, freely available open-source software framework, allowing readers to try out the techniques after reading the explanations. The book first offers a brief introduction to the topic, covering big data mining, basic methodologies for mining data streams, and a simple example of MOA. More detailed discussions follow, with chapters on sketching techniques, change, classification, ensemble methods, regression, clustering, and frequent pattern mining. Most of these chapters include exercises, an MOA-based lab session, or both. Finally, the book discusses the MOA software, covering the MOA graphical user interface, the command line, use of its API, and the development of new methods within MOA. The book will be an essential reference for readers who want to use data stream mining as a tool, researchers in innovation or data stream mining, and programmers who want to create new algorithms for MOA.

In the face of conflicting claims about some treatments, behaviors, and policies, the question arises: What is the most scientifically rigorous way to draw conclusions about cause and effect in the study of humans? In this introduction to causal inference, Paul Rosenbaum explains key concepts and methods through real-world examples.

A comprehensive introduction to Support Vector Machines and related kernel methods. In the 1990s, a new type of learning algorithm was developed, based on results from statistical learning theory: the Support Vector Machine (SVM). This gave rise to a new class of theoretically elegant learning machines that use a central concept of SVMs—-kernels—for a number of learning tasks. Kernel machines provide a modular framework that can be adapted to different tasks and domains by the choice of the kernel function and the base algorithm. They are replacing neural networks in a variety of fields, including engineering, information retrieval, and bioinformatics. Learning with Kernels provides an introduction to SVMs and related kernel methods. Although the book begins with the basics, it also includes the latest research. It provides all of the concepts necessary to enable a reader equipped with some basic mathematical knowledge to enter the world of machine learning using theoretically well-founded yet easy-to-use kernel algorithms and to understand and apply the powerful algorithms that have been developed over the last few years.

Causality in a Social World introduces innovative new statistical research and strategies for investigating moderated intervention effects, mediated intervention effects, and spill-over effects using experimental or quasi-experimental data. The book uses potential outcomes to

define causal effects, explains and evaluates identification assumptions using application examples, and compares innovative statistical strategies with conventional analysis methods. Whilst highlighting the crucial role of good research design and the evaluation of assumptions required for identifying causal effects in the context of each application, the author demonstrates that improved statistical procedures will greatly enhance the empirical study of causal relationship theory. Applications focus on interventions designed to improve outcomes for participants who are embedded in social settings, including families, classrooms, schools, neighbourhoods, and workplaces.
The Oxford Handbook of Causal Reasoning offers a state-of-the-art review of one of our most central cognitive competencies, which has for a long time been neglected in cognitive psychology. This Handbook provides introductions of competing theories of causal reasoning, and discusses its role in various cognitive functions and domains.
A comprehensive review of an area of machine learning that deals with the use of unlabeled data in classification problems: state-of-the-art algorithms, a taxonomy of the field, applications, benchmark experiments, and directions for future research. In the field of machine learning, semi-supervised learning (SSL) occupies the middle ground, between supervised learning (in which all training examples are labeled) and unsupervised learning (in which no label data are given). Interest in SSL has increased in recent years, particularly because of application domains in which unlabeled data are plentiful, such as images, text, and bioinformatics. This first comprehensive overview of SSL presents state-of-the-art algorithms, a taxonomy of the field, selected applications, benchmark experiments, and perspectives on ongoing and future research.Semi-Supervised Learning first presents the key assumptions and ideas underlying the field: smoothness, cluster or low-density separation, manifold structure, and transduction. The core of the book is the presentation of SSL methods, organized according to algorithmic strategies. After an examination of generative models, the book describes algorithms that implement the low-density separation assumption, graph-based methods, and algorithms that perform two-step learning. The book then discusses SSL applications and offers guidelines for SSL practitioners by analyzing the results of extensive benchmark experiments. Finally, the book looks at interesting directions for SSL research. The book closes with a discussion of the relationship between semi-supervised learning and transduction.
Recent arguments concerning the nature of causation in evolutionary theory, now often known as the debate between the 'causalist' and 'statisticalist' positions, have involved answers to a variety of independent questions – definitions of key evolutionary concepts like natural selection, fitness, and genetic drift; causation in multi-level systems; or the nature of evolutionary explanations, among others. This Element offers a way to disentangle one set of these questions surrounding the causal structure of natural selection. Doing so allows us to clearly reconstruct the approach that some of these major competing interpretations of evolutionary theory have to this causal structure, highlighting particular features of philosophical interest within each. Further, those features concern problems not exclusive to the philosophy of biology. Connections between them and, in two case studies, contemporary metaphysics and philosophy of physics demonstrate the potential value of broader collaboration in the understanding of evolution.
A graphical model is a statistical model that is represented by a graph. The factorization properties underlying graphical models facilitate tractable computation with multivariate distributions, making the models a valuable tool with a plethora of applications. Furthermore, directed graphical models allow intuitive causal interpretations and have become a cornerstone for causal inference. While there exist a number of excellent books on graphical models, the field has grown so much that individual authors can hardly cover its entire scope. Moreover, the field is interdisciplinary by nature. Through chapters by leading researchers from different

areas, this handbook provides a broad and accessible overview of the state of the art. Key features: * Contributions by leading researchers from a range of disciplines * Structured in five parts, covering foundations, computational aspects, statistical inference, causal inference, and applications * Balanced coverage of concepts, theory, methods, examples, and applications * Chapters can be read mostly independently, while cross-references highlight connections The handbook is targeted at a wide audience, including graduate students, applied researchers, and experts in graphical models.

This innovative research design text will help you make informed choices when carrying out your research project. Covering both qualitative and quantitative approaches, and with examples drawn from a wide range of social science disciplines, the authors explain what is at stake when choosing a research design, and discuss the trade-offs that researchers have to make when considering issues such as: - causality - categories and classification - heterogeneity - interdependence - time This book will appeal to students and researchers looking for an in-depth understanding of research design issues to help them design their projects in a thoughtful and responsible way.

This book honours the outstanding contributions of Vladimir Vapnik, a rare example of a scientist for whom the following statements hold true simultaneously: his work led to the inception of a new field of research, the theory of statistical learning and empirical inference; he has lived to see the field blossom; and he is still as active as ever. He started analyzing learning algorithms in the 1960s and he invented the first version of the generalized portrait algorithm. He later developed one of the most successful methods in machine learning, the support vector machine (SVM) – more than just an algorithm, this was a new approach to learning problems, pioneering the use of functional analysis and convex optimization in machine learning. Part I of this book contains three chapters describing and witnessing some of Vladimir Vapnik's contributions to science. In the first chapter, Léon Bottou discusses the seminal paper published in 1968 by Vapnik and Chervonenkis that lay the foundations of statistical learning theory, and the second chapter is an English-language translation of that original paper. In the third chapter, Alexey Chervonenkis presents a first-hand account of the early history of SVMs and valuable insights into the first steps in the development of the SVM in the framework of the generalised portrait method. The remaining chapters, by leading scientists in domains such as statistics, theoretical computer science, and mathematics, address substantial topics in the theory and practice of statistical learning theory, including SVMs and other kernel-based methods, boosting, PAC-Bayesian theory, online and transductive learning, loss functions, learnable function classes, notions of complexity for function classes, multitask learning, and hypothesis selection. These contributions include historical and context notes, short surveys, and comments on future research directions. This book will be of interest to researchers, engineers, and graduate students engaged with all aspects of statistical learning.

An accessible, contemporary introduction to the methods for determining cause and effect in the social sciences "Causation versus correlation has been the basis of arguments--economic and otherwise--since the beginning of time.

Causal Inference: The Mixtape uses legit real-world examples that I found genuinely thought-provoking. It's rare that a book prompts readers to expand their outlook; this one did for me."--Marvin Young (Young MC) Causal inference encompasses the tools that allow social scientists to determine what causes what. In a messy world, causal inference is what helps establish the causes and effects of the actions being studied--for example, the impact (or lack thereof) of increases in the minimum wage on employment, the effects of early childhood education on incarceration later in life, or the influence on economic growth of introducing malaria nets in developing regions. Scott Cunningham introduces students and practitioners to the methods necessary to arrive at meaningful answers to the questions of causation, using a range of modeling techniques and coding instructions for both the R and the Stata programming languages.

How does an algebraic geometer studying secant varieties further the understanding of hypothesis tests in statistics? Why would a statistician working on factor analysis raise open problems about determinantal varieties? Connections of this type are at the heart of the new field of "algebraic statistics". In this field, mathematicians and statisticians come together to solve statistical inference problems using concepts from algebraic geometry as well as related computational and combinatorial techniques. The goal of these lectures is to introduce newcomers from the different camps to algebraic statistics. The introduction will be centered around the following three observations: many important statistical models correspond to algebraic or semi-algebraic sets of parameters; the geometry of these parameter spaces determines the behaviour of widely used statistical inference procedures; computational algebraic geometry can be used to study parameter spaces and other features of statistical models.

Sections include: experiments and generalised causal inference; statistical conclusion validity and internal validity; construct validity and external validity; quasi-experimental designs that either lack a control group or lack pretest observations on the outcome; quasi-experimental designs that use both control groups and pretests; quasi-experiments: interrupted time-series designs; regresssion discontinuity designs; randomised experiments: rationale, designs, and conditions conducive to doing them; practical problems 1: ethics, participation recruitment and random assignment; practical problems 2: treatment implementation and attrition; generalised causal inference: a grounded theory; generalised causal inference: methods for single studies; generalised causal inference: methods for multiple studies; a critical assessment of our assumptions.

Human beings are active agents who can think. To understand how thought serves action requires understanding how people conceive of the relation between cause and effect, between action and outcome. In cognitive terms, how do people construct and reason with the causal models we use to represent our world? A revolution is occurring in how statisticians, philosophers, and computer scientists answer this question. Those fields have ushered in new insights about causal models by thinking about how to represent causal structure mathematically, in a framework that uses graphs and probability theory to develop what are called causal Bayesian networks. The framework starts with the idea that the purpose of causal structure is to understand and predict the effects of intervention. How does intervening on one thing affect other things? This is not a question merely about probability (or logic), but about action. The framework offers a new understanding of mind: Thought is about the effects of intervention and cognition is thus intimately tied to actions that take place either in the actual physical world or in imagination, in counterfactual worlds. The book offers a conceptual introduction to the key mathematical ideas, presenting them in a non-technical way, focusing on the intuitions rather than the theorems. It tries to show why the ideas are important to understanding how people explain things and why thinking not only about the world as it is but the world as it could be is so central to human action. The book reviews the role of causality, causal models, and intervention in the basic human cognitive functions: decision making, reasoning, judgment, categorization, inductive inference, language, and learning. In short, the book offers a discussion about how people think, talk, learn, and explain things in causal terms, in terms of action and manipulation.

A concise and self-contained introduction to causal inference, increasingly important in data science and machine learning. The mathematization of causality is a relatively recent development, and has become increasingly important in data science and machine learning. This book offers a self-contained and concise introduction to causal models and how to learn them from data. After explaining the need for causal models and discussing some of the principles underlying causal inference, the book teaches readers how to use causal models: how to compute intervention distributions, how to infer causal models from observational and interventional data, and how causal ideas could be exploited for classical machine learning problems. All of these topics are discussed first in terms of two variables and then in the more general multivariate case. The bivariate case turns out to be a particularly hard problem for causal learning because there are no conditional independences as used by classical methods for solving multivariate cases. The authors consider analyzing statistical asymmetries between cause and effect to be highly instructive, and they report on their decade of intensive research into this problem. The book is accessible to readers with a background in machine learning or statistics, and can be used in graduate courses or as a reference for researchers. The text includes code snippets that can be copied and pasted, exercises, and an appendix with a summary of the most important technical concepts.

A new approach for defining causality and such related notions as degree of responsibility, degrees of blame, and causal explanation. Causality plays a central role in the way people structure the world; we constantly seek causal explanations for our observations. But what does it even mean that an event C “actually caused” event E? The problem of defining actual causation goes beyond mere philosophical speculation. For example, in many legal arguments, it is precisely what needs to be established in order to determine responsibility. The philosophy literature has been struggling with the problem of defining causality since Hume. In this book, Joseph Halpern explores actual causality, and such related notions as degree of responsibility, degree of blame, and causal explanation. The goal is to arrive at a definition of causality that matches our natural language usage and is helpful, for example, to a jury deciding a legal case, a programmer looking for the line of code that cause some software to fail, or an economist trying to determine whether austerity caused a subsequent depression. Halpern applies and expands an approach to causality that he and Judea Pearl developed, based on structural equations. He carefully formulates a definition of causality, and building on this, defines degree of responsibility, degree of blame, and causal explanation. He concludes by discussing how these ideas can be applied to such practical problems as accountability and program verification. Technical details are generally confined to the final section of each chapter and can be skipped by non-mathematical readers.

This book compiles and presents new developments in statistical causal inference. The accompanying data and computer programs are publicly available so readers may replicate the model development and data analysis presented in each chapter. In this way, methodology is taught so that readers may implement it directly. The book brings together experts engaged in causal inference research to present and discuss recent issues in causal inference methodological development. This is also a timely look at causal inference applied to scenarios that range from clinical trials to mediation and public health research more broadly. In an academic setting, this book will serve as a reference and guide to a course in causal inference at the graduate level (Master's or Doctorate). It is particularly relevant for students pursuing degrees in statistics, biostatistics, and computational biology. Researchers and data analysts in public health and biomedical research will also find this book to be an important reference.