In this chapter a number of methods that make use of both graphical procedures and statistical methods are described and illustrated. Methods such as trend lines, median lines, and graphical representations of background variability are described and illustrated. Also described is a method especially designed for use by practitioners. Some of these methods make use of procedures adopted from statistical process control methodology. A recommended analysis procedure is described that makes use of both median-based and trend-line based methods for representing background variability.

The field of molecular evolution has experienced explosive growth in recent years due to the rapid accumulation of genetic sequence data, continuous improvements to computer hardware and software, and the development of sophisticated analytical methods. The increasing availability of large genomic data sets requires powerful statistical methods to analyse and interpret them, generating both computational and conceptual challenges for the field. This book provides a comprehensive coverage of modern statistical and computational methods used in molecular evolutionary analysis, such as maximum likelihood and Bayesian statistics. It describes the models, methods and algorithms that are most useful for analysing the ever-increasing supply of molecular sequence data, with a view to furthering our
understanding of the evolution of genes and genomes. The book emphasizes essential concepts rather than mathematical proofs. It includes detailed derivations and implementation details, as well as numerous illustrations, worked examples, and exercises.

An Introduction to Quantitative Methods Especially Relevant for Public Mental Health


in Public Mental Health

Published in print: 2012 Published Online: September 2012
DOI: 10.1093/acprof:oso/9780195390445.003.0005
Item type: chapter

This chapter provides a brief introduction to some of the epidemiologic and statistical methods for and challenges to gathering and analyzing the data that underlie the research presented in this volume and in the field of public mental health as a whole. It focuses on some of the data and methodological complexities particularly common in public mental health research. It explores three fundamental types of questions relevant to public mental health: estimating rates of disorders in a population across people, place, and time; examining risk and protective factors associated with particular disorders; and exploring interventions to prevent disorders or to treat them once they emerge.

The Mechanical, the Historical, and the Statistical

John J. McDermott (ed.)

in The Basic Writings of Josiah Royce, Volume II: Logic, Loyalty, and Community

Published in print: 2005 Published Online: May 2019
DOI: 10.5422/fordham/97808232324845.003.0003
Item type: chapter

This chapter begins by describing vitalism and materialism. The name “vitalism” is often given to those doctrines which have used the hypothesis that the phenomena of living organisms are due to some process which is essentially identical in its nature with the process exemplified by people's conscious voluntary activities. On the other hand, some things and events in the natural world—such as the recurrent movements of the heavenly bodies and the processes which attend the workings of machines—seem to be in many respects essentially different from the processes which result from people's plans, choices,
and voluntary deeds. What is called a “mechanical theory of nature” or “materialism” undertakes to account for the vital processes, for the activities of organisms. The chapter then considers the three classifications of scientific methods: the historical, the mechanical, and the statistical.

Statistics, Data Mining, and Machine Learning in Astronomy
Željko Ivezic, Andrew J. Connolly, Jacob T VanderPlas, and Alexander Gray

As telescopes, detectors, and computers grow ever more powerful, the volume of data at the disposal of astronomers and astrophysicists will enter the petabyte domain, providing accurate measurements for billions of celestial objects. This book provides a comprehensive and accessible introduction to the cutting-edge statistical methods needed to efficiently analyze complex data sets from astronomical surveys such as the Panoramic Survey Telescope and Rapid Response System, the Dark Energy Survey, and the upcoming Large Synoptic Survey Telescope. It serves as a practical handbook for graduate students and advanced undergraduates in physics and astronomy, and as an indispensable reference for researchers. The book presents a wealth of practical analysis problems, evaluates techniques for solving them, and explains how to use various approaches for different types and sizes of data sets. For all applications described in the book, Python code and example data sets are provided. The supporting data sets have been carefully selected from contemporary astronomical surveys (for example, the Sloan Digital Sky Survey) and are easy to download and use. The accompanying Python code is publicly available, well documented, and follows uniform coding standards. Together, the data sets and code enable readers to reproduce all the figures and examples, evaluate the methods, and adapt them to their own fields of interest.

Functional Divergence after Gene Duplication: Statistical Modeling
Xun Gu

in Statistical Theory and Methods for Evolutionary Genomics
Many organisms have undergone genome-wide or local chromosome duplication events during their evolution. As a result, many genes are represented as several paralogs in the genome with related but distinct functions (gene families). Since gene duplication is thought to have provided the raw materials for functional innovations, it is desirable to identify amino acid sites that are responsible for functional divergence from the sequence analysis of a gene family. A series of statistical models has been developed, based on the principle that functional divergence between duplicate genes is highly correlated with the change of evolutionary rate after the gene duplication. This chapter discusses these statistical and computational methods. These include the Poisson-gamma model for protein sequence evolution, the Markov chain model for type-I functional divergence, and statistical methods for type-II functional divergence.

Statistical Methods for Diffusion Data
William Hoppitt and Kevin N. Laland

in Social Learning: An Introduction to Mechanisms, Methods, and Models
Published in print: 2013 Published Online: October 2017
DOI: 10.23943/princeton/9780691150703.003.0005
princeton/9780691150703.003.0005
Item type: chapter

This chapter describes statistical methods for inferring and quantifying social transmission in groups of animals in the wild, or in “captive” groups of animals in naturalistic social environments. In particular, it considers techniques for analyzing time-structured data on the occurrence of a particular behavior pattern, or behavioral trait, in one or more groups. For the most part, the focus is on cases where a novel trait spreads through one or more groups. Following standard terminology in the field of social learning, the spread of a trait through a group is referred to as a diffusion, and the resulting data as diffusion data. The methods include diffusion curve analysis and network-based diffusion analysis. For the latter, inclusion of individual-level variables is taken into account, along with model selection and inference, modeling of multiple diffusions, choosing a social network, and “untransmitted” social effects. The chapter also examines the spatial spread of a behavioral trait.
Causal Mechanisms and Process Tracing
Gary Goertz and James Mahoney

in A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences
Published in print: 2012 Published Online: October 2017
Item type: chapter

This chapter examines how the qualitative and quantitative research traditions empirically assess theories about mechanisms when making causal inferences. In the qualitative paradigm, researchers carry out this assessment by attempting to observe causal mechanisms through process tracing and through the analysis of causal-process observations. In the qualitative paradigm, the within-case analysis of specific cases is combined with the effort to observe mechanisms. By contrast, statistical methods are not designed to observe mechanisms within particular cases. The chapter considers the importance of mechanisms in causal inference as well as the use of process tracing in multimethod vs. qualitative research. It shows that multimethod research, which integrates regression and case study analysis, is increasingly regarded as a best practice.

Stochastic Population Dynamics in Ecology and Conservation
Russell Lande, Steinar Engen, and Bernt-Erik Saether
Published in print: 2003 Published Online: April 2010
Item type: book

All populations fluctuate stochastically, creating a risk of extinction that does not exist in deterministic models, with fundamental consequences for both pure and applied ecology. This book provides an introduction to stochastic population dynamics, combining classical background material with a variety of modern approaches, including previously unpublished results by the authors, illustrated with examples from bird and mammal populations, and insect communities. Demographic and environmental stochasticity are introduced with statistical methods for estimating them from field data. The long-run growth rate of a population is explained and extended to include age structure with both demographic and environmental stochasticity. Diffusion approximations facilitate the analysis of extinction dynamics and the duration of the final decline. Methods are developed for estimating delayed density dependence from population time series using life history data. Metapopulation
viability and the spatial scale of population fluctuations and extinction risk are analyzed. Stochastic dynamics and statistical uncertainty in population parameters are incorporated in Population Viability Analysis and strategies for sustainable harvesting. Statistics of species diversity measures and species abundance distributions are described, with implications for rapid assessments of biodiversity, and methods are developed for partitioning species diversity into additive components. Analysis of the stochastic dynamics of a tropical butterfly community in space and time indicates that most of the variance in the species abundance distribution is due to ecological heterogeneity among species, so that real communities are far from neutral.

Introduction

William Hoppitt and Kevin N. Laland

in Social Learning: An Introduction to Mechanisms, Methods, and Models

This book explores social learning concepts and methods as well as new developments in the field. These methods include: experimental and statistical methods that allow researchers to categorize cases of social learning according to their underlying psychological processes and learning mechanisms; neuroscientific methods for identifying the brain structures, neural circuitry, and physiological processes underlying both social learning and social influences on decision making; and mathematical methods for predicting the pattern of diffusion of novel learned innovations, and for modeling cultural evolution and gene-culture coevolution. This introductory chapter presents some definitions such as “social learning,” “imitation,” “innovation,” and “social transmission” and explains why social learning is an important field of study. It also provides an overview of the chapters that follow.

Combining results from several studies: systematic review and meta-analysis

J. Mark Elwood

in Critical Appraisal of Epidemiological Studies and Clinical Trials
This chapter is divided into two parts. The first part discusses systematic reviews and the role of meta-analysis. The second part presents the most widely used statistical methods for meta-analysis and related issues such as exploring differences between studies. Self-test questions are provided at the end of the chapter.

**The Production of Knowledge**

William H. Starbuck

in *The Production of Knowledge: The Challenge of Social Science Research*

Published in print: 2006 Published Online: October 2011
Item type: chapter

Because social science research has a low utilitarian value, little attention is given to such research and social scientists are probably receiving a lot less respect than they deserve. Higher standards for reliability should then be set for social science research so that more respect would be given to the researchers and the findings would have more value for society. This chapter presents some research tactics that will hopefully improve the results of social science research or at least filter unnecessary noise. The chapter encourages researchers to challenge their own thinking to be able to demonstrate the validity of their knowledge through observation and displacing situations from equilibria. Also, the chapter introduces statistical methods that would yield the production of reliable knowledge. Lastly, researchers should be able to create consensus about what they know.

**Event Time Analysis**

Ezra Susser, Sharon Schwartz, Alfredo Morabia, and Evelyn J. Bromet

in *Psychiatric Epidemiology: Searching for the Causes of Mental Disorders*

Published in print: 2006 Published Online: September 2009
Item type: chapter

This chapter describes statistical methods for taking account of unequal attrition, that is, different follow-up times across exposed and unexposed groups in a cohort study. It considers methods for analyzing time to an event as an outcome in its own right. The chapter defines the outcome as the time until disease onset. From this perspective, a nondiseased subject in the study is not a control as in a case-control study, but someone who has not yet developed the disease. To analyze disease
incidence data from this perspective, methods that come under the interchangeable headings of survival analysis, time-to-event analysis, or failure-time analysis are used.

Social Learning Strategies
William Hoppitt and Kevin N. Laland

in Social Learning: An Introduction to Mechanisms, Methods, and Models
Published in print: 2013 Published Online: October 2017
DOI: 10.23943/princeton/9780691150703.003.0008
Item type: chapter

This chapter focuses on social learning strategies—functional rules specifying what, when, and who to copy. There are many plausible social learning strategies. Individuals might disproportionately copy when asocial learning would be difficult or costly, when they are uncertain of what to do, when the environment changes, when established behavior proves unproductive, and so forth. Likewise, animals might preferentially copy the dominant individual, the most successful individual, or a close relative. This chapter presents evidence for some of the better-studied learning heuristics and describes statistical procedures for identifying which social learning strategies are being deployed in a data set. It examines “who” strategies, which cover frequency-dependent biases, success biases, and kin and age biases, as well as “what” strategies, random copying, and statistical methods for detecting social learning strategies. Finally, it evaluates meta-strategies, best strategies, and hierarchical control.

Measuring Catholic Faith in Nineteenth- and Early Twentieth-Century Northeast China
Ji Li

in In God’s Empire: French Missionaries and the Modern World
Published in print: 2012 Published Online: September 2012
DOI: 10.1093/acprof:oso/9780195396447.003.0008
Item type: chapter

This chapter explores how missionaries, starting in the mid-nineteenth century, developed a method for assessing the degree of faith of Christians in communities in Manchuria. Far from being a passing experiment, missionaries’ use of statistical methods and quantitative language to measure the religiosity of local converts became a central concern of the Church, as important as translating texts to and from
Latin, French, and Chinese. In creating this method, missionaries calculated the impact of their efforts and defined what it meant to be a “good” Christian.

Epidemiological methods: a view from the Americas
Eduardo L. Franco, Salaheddin M. Mahmud, and Andrew G. Dean

in The Development of Modern Epidemiology: Personal reports from those who were there

Published in print: 2007 Published Online: September 2009
Item type: chapter

This chapter discusses the history of epidemiological methods. It provides a personal account of how the three authors view the evolution of study designs, statistical methodology, and the development of epidemiological computer software that aided progress in the field. Whenever appropriate, emphasis is given to the view from the Americas. However, in the interest of providing a seamless historical perspective, key scientific developments from other continents are also included.

Statistics and Scientific Method
Peter J. Diggle and Amanda G. Chetwynd

Published in print: 2011 Published Online: December 2013
Item type: book

An antidote to technique-oriented service courses, this book studiously avoids the recipe-book style and keeps algebraic details of specific statistical methods to the minimum extent necessary to understand the underlying concepts. Instead, it aims to give the reader a clear understanding of how core statistical ideas of experimental design, modelling, and data analysis are integral to the scientific method. Aimed primarily towards a range of scientific disciplines (albeit with a bias towards the biological, environmental, and health sciences), this book assumes some maturity of understanding of scientific method, but does not require any prior knowledge of statistics, or any mathematical knowledge beyond basic algebra and a willingness to come to terms with mathematical notation. Any statistical analysis of a realistically sized data-set requires the use of specially written computer software. An Appendix introduces the reader to our open-source software of choice. All of the material in the book can be understood without using either R or any other computer software.
Intergenerational data necessarily reflect the time and place that the different generations of participants were living in. This chapter aims first to introduce simple concepts to provide an understanding of the founding assumptions and principles, before moving on to more complex analytic methods. As the objectives of analyses may vary substantially across intergenerational studies, there is no easy guideline for analyses, except perhaps that some a priori clarity on the main associations of interest is crucial. Since parents and their offspring are genetically related, intergenerational studies are to some extent genetically informative even if no genotyping was performed. Much of the analyses are concerned with identifying or unravelling the relationship between outcomes and genetic and environmental factors. Ways of handling missing data as well as approaches to deal with non-paternity are also discussed. Illustrative examples are drawn from the two cohort studies.

Critical Appraisal of Epidemiological Studies and Clinical Trials

Mark Elwood

This book presents a logical system of critical appraisal, and facilitates studies to be evaluated and carried out effectively. This system emphasizes the central importance of cause and effect relationships. Its great strength is that it is applicable to a wide range of issues, and both to intervention trials and observational studies. This system unifies the often different approaches used in epidemiology, health services research, clinical trials, and evidence-based medicine, starting from a logical consideration of cause and effect. The book looks at the issues of study design, selection of subjects, bias, confounding, and the place of statistical methods. Systematic reviews, meta-analysis, and the applications of this logic to evidence-based medicine, knowledge-based health care, and health practice and policy are discussed. Current and often controversial examples are used, including screening for
prostate cancer, publication bias in psychiatry, public health issues in developing countries, and conflicts between observational studies and randomized trials. Statistical issues are explained clearly without complex mathematics, and the most useful methods are summarized in the appendix. The final chapters give six applications of the critical appraisal of major studies: randomized trials of medical treatment and prevention, a prospective and a retrospective cohort study, a small matched case-control study, and a large case-control study. In these chapters, sections of the original papers are reproduced and the original studies placed in context by a summary of current developments.

Family matters
Deborah A. Lawlor and Gita D. Mishra (eds)

Family based studies, including intergenerational, sibling and twin studies, are increasingly used to explore life course epidemiology. However, understanding the underlying assumptions of these studies and hence the inferences that can be drawn from them is complex. Further, there are issues relating to study design and the statistical analysis of family-based studies that are not well understood. This book provides in one volume the knowledge and skills required to design, analyse, and correctly interpret family based studies. The book is divided into four sections that cover the theoretical underpinning of using family based studies in life course epidemiology; practical issues of data collection, storage, and collaborative use (including a chapters on conducting such studies in low and middle income countries and the use of information provided by other family members); statistical analysis and correct interpretation of results from family based studies and a section that illustrates the use of these studies in life course epidemiology, with examples from three areas — cardiovascular disease, mental health and wellbeing, and reproductive health.