Contents

Foreword

This comprehensive book takes the reader from the underpinnings of statistical inference through to cutting-edge modern analytical techniques. Along the way, the authors explore graphical representations of data, a key component of any data analysis; standard procedures such as the *t*-test and tests for independence; and modern methods, including the bootstrap and empirical likelihood methods. The presentation focuses on practical applications interwoven with theoretical rationale, with an emphasis on how to carry out procedures and interpret the results. Numerous software examples (R and SAS) are provided, such that the readers should be able to reproduce plots and other analyses on their own. A wealth of examples from real data sets, web resources, supplemental notes, and plentiful references are provided, which round out the materials.

Nicole Lazar
Department of Statistics, University of Georgia
Athens, Georgia

Preface

SUMMARY: In all likelihood, the universe of statistical science is relatively or privately infinite, and it is expanding in a similar manner as our universe, satisfying the property of a science that is alive. Hence, it would be an absolute impossibility to include everything in one book written by humans, even around a statistical topic such as statistical testing strategies. One of our major goals is to provide the readers a road map of where and what to look for, corresponding to their interests, while steering them away from improper or suboptimal approaches.

Modern studies in health sciences require formal comparisons and evaluations based on correct and appropriate statistical decision-making mechanisms. In many scenarios, the statistical design and analysis of a health-related study is carried forth under strict federal guidelines and review; for example, the Food and Drug Administration (FDA) reviews and approves the safety and efficacy of new drugs and medical devices based on rigorous statistical approaches. The cost of these types of investigations is enormous and time-consuming. Thus, the use of cost-effective and appropriate statistical methods is a high priority. In addition to federal reviews, scientific journals have become much more rigorous as well, in terms of making sure an appropriate analysis plan was applied, usually employing statisticians as reviewers.

The primary objective of this book is to provide a compendium of statistical approaches for decision-making policies ranging from graphical methods and classical procedures through computational-intensive bootstrap strategies to newly advanced empirical likelihood techniques. Historically, many of the approaches that we will present in this book were not feasible due to the computational requirements. In the past, this meant relying on approximations that may or may not be accurate in the finite sample setting. However, in the modern age, we are generally no longer constrained by computational issues and have a greater flexibility in terms of the statistical approaches that we may employ to data analysis problems. In some sense, the computational power of modern computers has spurred a renaissance in new and modern statistical methods, which is the thrust of what we wish to accomplish. These methodologies may be applied to various problems encountered in medical and epidemiological studies, including clinical trials.

Theoretical aspects of various decision-making procedures, including those based on parametric and nonparametric likelihood methods, have been extensively discussed in the literature. We introduce correct and efficient testing mechanisms with the aim of making these techniques assessable to a wide variety of data analysts at different skill levels by providing the accompanying software routines. In terms of the book's content, we cannot cover every statistical testing scenario and strategy. For example, we do not include tests based on survival data or topics such as regression models. In general, the techniques covered in this book span robust statistical methods to more computationally intensive approaches.

Our intent is that this book provides material that will bridge the gap between highly relevant theoretical statistical methods and practical procedures applied to the planning and analysis of health-related experiments. The theoretical underpinnings of the methods presented in this book are introduced with a substantial amount of chapter notes to point the readers to additional references. We work from a theoretical framework and move toward the practical application for each concept, providing real-world examples that are made easily implementable by providing the corresponding statistical software code. We also show the basic ingredients and methods for constructing correct and powerful statistical decision-making processes to be adapted toward complex statistical applications.

We start the book by laying the foundation for properly framing health-related experiments in terms of formal statistical decision rules. We also include novel theoretical and applied results, which have been published in high-impact biostatistical and statistical journals, including those developed by the authors. In order to help transfer statistical theory into practice, we provide easily accessible software routines for a majority of the methods described in the book based on both the R and SAS statistical software packages. Descriptions of the R and SAS languages at an introductory level are also presented in our book.

Generally speaking, we try to balance presenting the materials in a manner that is useful to medical researchers who have some training in statistics and statistical computing and need only a reminder of some of the basic statistical tools and concepts with the interests of research-oriented statisticians who are concerned with the theoretical aspects of the methods we present. In the book, the reader will find new theoretical methods, open problems, and new procedures across a variety of topics for their scholarly investigations. We present results that are novel to the current set of books on the market and results that are even new with respect to the recent scientific literature. Our aim is to draw the attention of theoretical statisticians and practitioners in epidemiology and clinical research to the necessity of new developments, extensions, and investigations related to statistical testing and its applications. In the context of focusing this book on applied work, we emphasize the following topics:

- Correctly formulating statistical hypotheses with respect to aims of medical and epidemiological studies
- Constructing and providing statistical decision-making test rules corresponding to practical experiments
- Using simple but efficient graphical methods, including the receiver operating characteristic (ROC) curve analysis and Kendall-type plots
- Using basic test procedures and their components in practical decision-making problems
- Employing parametric and nonparametric likelihood testing techniques in applied research
- Understanding the basic properties of likelihood ratio-type tests in parametric and nonparametric settings
- Using Bayesian concepts to develop decision-making procedures
- Applying goodness-of-fit tests and tests for independence
- Understanding the concepts of retrospective and sequential decision-making mechanisms
- Solving multiple testing problems in clinical experiments
- Defining and solving change point detection problems
- Comparing statistical tests and evaluating their properties
- Understanding and employing bootstrap and permutation methods
- Implementing the statistical software at a beginning level
- Learning modern techniques and their applications presented in the recent statistical literature

For the more research-oriented statistician, our book describes important properties and new as well as classical theoretical results of testing statistical hypotheses. We introduce several methodological tools to evaluate different theoretical characteristics of statistical procedures, for example, Chapters 4, 6, and 7. We present higher-level theoretical developments relative to parametric and empirical likelihood functions and Bayesian approaches to statistical inference, which may be extended, modified, and investigated using asymptotic and finite sample size propositions. We lay the foundation for new dependence measures based on Kendall plots in terms of both summary statistical measures and graphical displays. New methods for combining biomarkers based on ROC methods are presented. Learning the book material, one can easily find interesting open problems. For example, in Chapter 1 we introduce the approach of expected p-values. The reader can recognize that the formal notation for expected p-values has a form of the area under the ROC curve (AUC). Then the problem to obtain the best combinations of different test statistics (Chapter 14) can be stated in a manner related to the maximization of AUCs, mentioned in Chapter 8. The introduced bootstrap-type techniques (e.g., Chapter 17) are very attractive to be employed in this context. In a similar mode, the areas under the Kendal plot (Chapter 10) can be considered. The software code mentioned in this book can certainly be improved, optimized, and extended.

This book describes statistical strategies for handling multiplicity issues arising in hypothesis testing across medical and epidemiological studies. We cover a variety of bootstrap methods from a practical point of view, making sure the limitations of the approach are understood in the finite sample setting, and develop strategies for scenarios when the most basic bootstrap approach may fail. We also introduce the concept of permutation—or so-called exact tests—which under certain conditions is a very desirable and robust testing strategy. We display traditional statistical testing techniques as well as present novel testing approaches. The book shows efficient methods for constructing powerful procedures to test for composite hypotheses based on data that can be subject to different issues related to problematic data, for example, data with observations containing measurement errors or limit of detection issues. We introduce several strategies to evaluate and compare treatment effects based on multiple biomarkers. Our book offers tools that help the reader to fully understand statistical approaches for decision making. This book points out comparisons between different statistical methods based on real medical data.

The organization of the material within the book is primarily based on the type of questions to be answered by inference procedures or according to the general type of mathematical derivation. We consider a wide class of decision-making mechanisms, starting from probability plotting techniques via the receiver operating characteristic curve analysis, which is widely applied in diagnostic studies, and moving to modern nonparametric likelihood approaches. The arrangement of chapters makes the book attractive to scientists who are new to the biostatistical research area, including those who may not have a strong statistical background. It can help attract statisticians interested in learning more about advanced topics. We explain step-by-step procedures and present a practical opportunity for trainees not only to acquaint themselves with statistics in the health sciences and understand statistical analysis in the medical literature, but also to be guided in the application of planning and analysis in their own studies.

The book is also intended for graduate students majoring in biostatistics, epidemiology, health-related sciences, or in a field where a statistics concentration is desirable, particularly for those who are interested in formal decision-making mechanisms. This book can be used as a textbook for a one- or two-semester advanced graduate course. The material in the book should be appropriate for use as both a text and a reference. We hope that the mathematical level and breadth of examples will recruit students and teachers not only from statistics and biostatistics, but also from a broad range of fields. We anticipate the book will induce the readers to learn a variety of new and current statistical techniques, and that by employing these new and efficient methods, they will change the practice of data analysis in clinical research.

A very important reason for writing this book was to refocus the scientist toward better understanding the underpinnings of appropriate statistical inference in a well-rounded fashion. It is our experience that applied statisticians or users often neglect the underlying postulates when implementing formal test procedures and interpreting their results. Consider, for example, the following telling quote:

> *P* values are widely used in science to test null hypotheses. For example, in a medical study looking at smoking and cancer, the null hypothesis could be that there is no link between the two. Many researchers interpret a lower *P* value as stronger evidence that the null hypothesis is false. Many also accept findings as 'significant' if the *P* value comes in at less than 0.05. But *P* values are slippery, and sometimes, significant *P* values vanish when experiments and statistical analyses are repeated. (Nature 506, 150–152, 2014.)

In this book, we provide the definition of a p-value and describe its formal role in decision-making mechanisms (e.g., see Chapter 1). Under the null hypothesis, p-values are uniformly [0,1] distributed random variables. At this rate, it is difficult to use p-values as the measurement. Oftentimes, practitioners misinterpret the notion of a p-value relative to a given statistical test. Without considering the statistical design of a given experiment and other information, careless interpretations of p-values can yield to absurd and unrealistic conclusions. This has resulted in different cases when practitioners try to avoid applications of statistical tools. For example, in 2015, the editors of Basic and Applied Social Psychology (*BASP*) announced that the journal would no longer publish papers containing p-values because the statistics were too often used to support low-quality research.

Finally, we would like to note that this book attempts to represent a part of our life that consists of mistakes, stereotypes, puzzles, and so forth—all that we love. Thus, our book cannot be perfect. We truly thank the reader for his or her participation in our life. We hope that the presented material can play a role as prior information for various research outputs.

Albert Vexler
Alan D. Hutson
Xiwei Chen

Authors

Albert Vexler obtained his PhD in statistics and probability theory from the Hebrew University of Jerusalem in 2003. His PhD advisors were Moshe Pollak, fellow of the American Statistical Association, and Marcy Bogen, professor of statistics at Hebrew University. He was a postdoctoral research fellow in the biometry and mathematical statistics branch at the National Institute of Child Health and Human Development (National Institutes of Health, Bethesda, Maryland). Currently, Dr. Vexler is a tenured associate professor at the Department of Biostatistics, The State University of New York at Buffalo. Dr. Vexler has authored and coauthored various publications that contribute to both the theoretical and applied aspects of statistics in medical research. Many of his papers and statistical software developments have appeared in statistical and biostatistical journals that have top-rated impact factors and are historically recognized as leading scientific journals, including *Biometrics*, *Biometrika*, *Journal of Statistical Software*, *American Statistician*, *Annals of Applied Statistics*, *Statistical Methods in Medical Research*, *Biostatistics*, *Journal of Computational Biology*, *Statistics in Medicine*, *Statistics and Computing*, *Computational Statistics and Data Analysis*, *Scandinavian Journal of Statistics*, *Biometrical Journal*, *Statistics in Biopharmaceutical Research*, *Stochastic Processes and Their Applications*, *Journal of Statistical Planning and Inference*, *Annals of the Institute of Statistical Mathematics*, *Canadian Journal of Statistics*, *Metrika*, *Statistics, Journal of Applied Statistics*, *Journal of Nonparametric Statistics*, *Communications in Statistics*, *Sequential Analysis*, *STATA Journal*, *American Journal of Epidemiology*, *Epidemiology*, *Paediatric and Perinatal Epidemiology*, *Academic Radiology*, *Journal of Clinical Endocrinology and Metabolism*, *Journal of Addiction Medicine*, and *Reproductive Toxicology and Human Reproduction*.

Dr. Vexler was awarded a National Institutes of Health grant to develop novel nonparametric data analysis and statistical methodology. His research interests include receiver operating characteristic curve analysis, measurement error, optimal designs, regression models, censored data, change point problems, sequential analysis, statistical epidemiology, Bayesian decision-making mechanisms, asymptotic methods of statistics, forecasting, sampling, optimal testing, nonparametric tests, empirical likelihoods, renewal theory, Tauberian theorems, time series, categorical analysis, multivariate analysis, multivariate testing of complex hypotheses, factor and principal component analysis, statistical biomarker evaluations, and best combinations of biomarkers. Dr. Vexler is the associate editor for *BMC Medical Research Methodology*, a peer-reviewed journal that considers articles on methodological approaches to healthcare research. In 2015, he was appointed the associate editor for *Biometrics*. These journals belong to the first cohort of academic literature related to the methodology of biostatistical and epidemiological research and clinical trials.

Alan D. Hutson received his BA in 1988 and MA in 1990 in statistics from the State University of New York (SUNY) at Buffalo. He then worked for Otsuka America Pharmaceuticals, Rockville, Maryland, for two years as a biostatistician. He then received his MA in 1993 and PhD in 1996 in statistics from the University of Rochester, Rochester, New York. His PhD advisor was Professor Govind Mudholkar, a world-renowned researcher in statistics and biostatistics. Dr. Hutson was hired as a biostatistician at the University of Florida, Gainesville, in 1996, as a research assistant professor and worked his way to a tenured associate professor. He had several roles at the University of Florida, including interim director of the Division of Biostatistics and director of the General Clinical Research Informatics Core. Dr. Hutson moved to SUNY at Buffalo in 2002 as an associate professor and chief of the Division of Biostatistics. He was the founding chair of the new Department of Biostatistics in 2003 and became a full professor in 2007. Dr. Hutson's accomplishments as chair included the implementation of several new undergraduate and graduate degree programs and a substantial growth in the size and quality of the department faculty and students. In 2005, Dr. Hutson also became chair of biostatistics (now biostatistics and bioinformatics) at Roswell Park Cancer Institute (RPCI), Buffalo, New York, was appointed professor of oncology, and was the director of the Core Cancer Center Biostatistics Core. He helped implement the new bioinformatics core at RPCI. He is currently chair of biostatistics at SUNY at Buffalo and chair of biostatistics and bioinformatics at RPCI. Dr. Hutson recently became the biostatistical, epidemiological, and research design director for SUNY's recently awarded National Institutes of Health–funded Clinical and Translational Research Award. Dr. Hutson is a fellow of the American Statistical Association. He is the associate editor of *Communications in Statistics*, the associate editor of the *Sri Lankan Journal of Applied Statistics*, and a New York State NYSTAR Distinguished Professor. Dr. Hutson has membership on several data safety and monitoring boards and has served on several high-level scientific review panels. He has more than 200 peer-reviewed publications. In 2013, Dr. Hutson was inducted into the Delta Omega Public Health Honor Society, Gamma Lambda Chapter. Dr. Hutson's methodological work focuses on nonparametric methods for biostatistical applications as they pertain to statistical functionals. He has several years of experience in the design and analysis of clinical trials.

Xiwei Chen obtained her bachelor's degree in mathematics from Peking University, Beijing, China, in 2008 and her MS degree in applied statistics from the Rochester Institute of Technology, Henrietta, New York, in 2010. In 2016, Ms. Chen obtained her PhD in biostatistics from the State University of New York at Buffalo. Her adviser was Dr. Albert Vexler. Between 2011 and 2015, she worked as a research assistant at Roswell Park Cancer Institute, Buffalo, New York. Currently, Ms. Chen is employed as a biostatistician at Johnson & Johnson Vision Care, Inc. Her areas of specialty are the empirical likelihood methods, the receiver operating characteristic curve methodology, and statistical diagnosis and its applications. Ms. Chen has authored or coauthored more than 10 papers and several book chapters in biostatistical areas concerning statistical approaches related to disease diagnoses. She is also very active as a reviewer for statistical journals.

1

Preliminaries: Welcome to the Statistical Inference Club: Some Basic Concepts in Experimental Decision Making

"I'm just one hundred and one, five months and a day."

"I can't believe that!" said Alice.

"Can't you?" the Queen said in a pitying tone. "Try again: draw a long breath, and shut your eyes."

Alice laughed. "There's no use trying," she said: "one can't believe impossible things."

"I daresay you haven't had much practice," said the Queen. "When I was your age, I always did it for half-an-hour a day. Why, sometimes I've believed as many as six impossible things before breakfast."

Lewis Carroll (Charles Lutwidge Dodgson), 1871*

1.1 Overview: Essential Elements of Defining Statistical Hypotheses and Constructing Statistical Tests

The majority of experiments in biomedicine and other health-related sciences involve mathematically formalized tests. The common goal in testing is to employ appropriate and efficient statistical procedures to make informed decisions based on data. Mathematical strategies for constructing formal decision rules play an important role in medical and epidemiological discovery, policy formulation, and clinical practice. In order to draw conclusions about populations on the basis of samples from populations, clinical experiments commonly require the application of the mathematical statistics discipline. It is often said that statistics is the language of applied science.

The aim of the scientific method in decision theory is to simultaneously maximize quantified gains and minimize losses when reaching a conclusion. For example, clinical experiments can have stated goals of maximizing factors (gains), such as the accuracy of diagnosis of a medical condition, faster healing, and greater patient satisfaction, while minimizing factors (losses), such as misdiagnoses, human resource expenditures, employed efforts, duration of screening for a disease, side effects, and costs of an experiment.

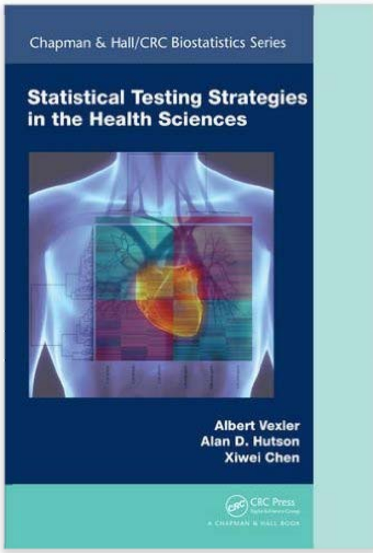
In order to correctly define statistical testing procedures and interpret the corresponding results, practitioners should research the nature of clinical data, experimental limitations, and instrumental sensitivities, as well as appropriate state objectives and their corresponding hypotheses.

* *Through the Looking-Glass, and What Alice Found There*, 1871. Lewis Carroll is a writer, mathematician, and logician.

Books › Science & Math › Biological Sciences

Statistical Testing Strategies in the Health Sciences (Chapman & Hall/CRC Biostatistics Series)

by Albert Vexler (Author), Alan D. Hutson (Author), Xiwei Chen (Author)



ISBN-13: 978-1498730815
ISBN-10: 1498730817
Why is ISBN important?

Add to List

Share

Hardcover
$102.72 - $108.93

Other Sellers
from $102.72

Buy used $102.72

Buy new **$108.93**

This title has not yet been released.
Ships from and sold by Amazon.com. Gift-wrap available.
FREE Prime Shipping once available

List Price: ~~$119.95~~ Save: $11.02 (9%)
Pre-order Price Guarantee.
3 New from $104.93

Qty: 1

Pre-order: Add to Cart

Turn on 1-Click ordering

Ship to:
Albert Vexler- Williamsville

More Buying Choices

4 used & new from $102.72

3 New from $104.93   1 Used from $102.72

See All Buying Options

Statistical Testing Strategies in the Health Sciences provides a compendium of statistical approaches for decision making, ranging from graphical methods and classical procedures through computationally intensive bootstrap strategies to advanced empirical likelihood techniques. It bridges the gap between theoretical statistical methods and practical procedures applied to the planning and analysis of health-related experiments.

The book is organized primarily based on the type of questions to be answered by inference procedures or

Read more

Special Offers and Product Promotions

- Pre-order Price Guarantee! Order now and if the Amazon.com price decreases between your order time and the end of the day of the release date, you'll receive the lowest price. Here's how (restrictions apply)
- Thank you for being a Prime member. Get a **$20.00 gift card instantly** upon approval for the **Amazon.com Store Card.**. Apply now.

Editorial Reviews

Review

"This comprehensive book takes the reader from the underpinnings of statistical inference through to cutting-edge modern analytical techniques. Along the way, the authors explore graphical representations of data, a key component of any data analysis; standard procedures such as the t-test and tests for independence; and modern methods, including the bootstrap and empirical likelihood method. The presentation focuses on practical applications interwoven with theoretical rationale, with an emphasis on how to carry out procedures and interpret the results. Numerous software examples (R and SAS) are provided, such that the readers should be able to reproduce plots and other analyses on their own. A wealth of examples from real data sets, web resources, supplemental notes, and plentiful references are provided, which round out the materials."
—From the Foreword by Nicole Lazar, Department of Statistics, University of Georgia

About the Author

Albert Vexler is a tenured associate professor in the Department of Biostatistics at the State University of New York (SUNY) at Buffalo. Dr. Vexler is the associate editor of *Biometrics* and *BMC Medical Research Methodology*. He is the author and coauthor of various publications that contribute to the theoretical and applied aspects of statistics in medical research. Many of his papers and statistical software developments have appeared in statistical and biostatistical journals that have top-rated impact factors and are historically recognized as leading scientific journals. Dr. Vexler was awarded a National Institutes of Health grant to develop novel nonparametric data analysis and statistical methodology. His research interests include receiver operating characteristic curve analysis, measurement error, optimal designs, regression models, censored data, change point problems, sequential analysis, statistical epidemiology, Bayesian decision-making mechanisms, asymptotic methods of statistics, forecasting, sampling, optimal testing, nonparametric tests, empirical likelihoods, renewal theory, Tauberian theorems, time series, categorical analysis, multivariate analysis, multivariate testing of complex hypotheses, factor and principal component analysis, statistical biomarker evaluations, and best combinations of biomarkers.

Alan D. Hutson is the chair of biostatistics and bioinformatics at Roswell Park Cancer Institute. He is also the biostatistical, epidemiological, and research design director for SUNY's National Institutes of Health–funded Clinical and Translational Research Award. Dr. Hutson is a fellow

Product Details

Series: Chapman & Hall/CRC Biostatistics Series
Hardcover: 703 pages
Publisher: Chapman and Hall/CRC (May 24, 2016)
Language: English
ISBN-10: 1498730817
ISBN-13: 978-1498730815
Shipping Weight: 1.7 pounds (View shipping rates and policies)
Amazon Best Sellers Rank: #1,227,333 in Books (See Top 100 in Books)
#292 in Books > Textbooks > Medicine & Health Sciences > Research > **Biostatistics**
#500 in Books > Medical Books > Basic Sciences > **Biostatistics**
#2302 in Books > Textbooks > Medicine & Health Sciences > **Reference**

Tell the Publisher!
I'd like to read this book on Kindle

Don't have a Kindle? Get your Kindle here, or download a **FREE** Kindle Reading App.

Customer Reviews

There are no customer reviews yet.

5 star
4 star
3 star
2 star
1 star

This item has not been released yet and is not eligible to be reviewed.