

## Book Review

Whitney Kirzinger<sup>1</sup>

**Lawrence Hubert and Howard Wainer.** *A Statistical Guide for the Ethically Perplexed*. 2013  
New York: CRC Press, ISBN 978-1-4398-7368-7, 565 pp., \$49.95.

This text draws attention to a topic that often seems to be neglected or at least carelessly regarded: the intersection of statistics and ethics. The work is modeled on a manuscript written by a medieval Jewish philosopher as an attempt to harmonize his philosophical views with Jewish law. Similarly, this text represents Hubert and Wainer's attempt to balance statistics and standards of ethical practice. The authors achieve this by providing interesting and relevant real-life examples covering a variety of topics, including the legal burden of proof, the use of statistics in the medical field, the ethics of data collection and data sleuthing, and the use (or misuse) of statistics in Supreme Court cases.

The work is composed of three sections. Part I is structured based on general statistical concepts. The authors discuss statistical tools, formulas, and theorems in an engaging and straightforward manner, using short stories and vignettes that even students beginning the study of statistics will enjoy reading. After introducing the subject in the first two chapters, Part I begins with Chapter 3, a discussion of probability theory and Bayes' theorem. This chapter is very useful for statistics students, as it describes Bayes' theorem several different ways: using mathematical formulas,  $2 \times 2$  contingency tables, and in simple nonscientific language. The reader will appreciate the clear examples that illustrate the practical implications of Bayes' theorem. Readers from all educational backgrounds will certainly be able to learn from the misunderstandings and misapplications of Bayes' theorem highlighted in this section. The emphasis is on the conclusion that even experienced statisticians can and historically have fallen prey to these common mistakes. The authors underscore one's duty to think critically about generating, interpreting, and reporting probabilities. Many examples of misused statistics are taken from well-known legal cases, with one interesting example being a problematic conditional probability used by a Harvard law professor who advised the O.J. Simpson defense team. The chapter on probability theory also elaborates on sensitivity, specificity, and positive predictive value in the context of breast cancer screening.

Chapter 4 focuses on application areas touching on concepts such as causation, relative risk, odds ratios, cohort studies and again provides great context and easy reading for the beginning statistician. The chapter also reviews simple experiments and sample space in the context of engaging topics such as spread betting, parimutuel betting, gaming, risk, and point shaving in college basketball. The chapter concludes by examining, with compelling

<sup>1</sup> Data Analysis and Quality Assurance Branch, Division of Health Interview Statistics, National Center for Health Statistics, 3311 Toledo Road, Room 2329, Hyattsville, Maryland 20782. Email: [wkirzinger@cdc.gov](mailto:wkirzinger@cdc.gov)

attention, the consequences of framing data used to make decisions, the role psychology of risk behavior can play in decision making, and the importance of being able to assess the quality of the information used when making decisions.

Next, Chapter 5 is a cautionary tale on the topic of correlation, as it specifically focuses on correlational fallacies. The first warning is against providing a correlation value without showing the associated scatterplot. Several arguments are given to demonstrate the need for the graphical representation of correlation data. The authors effectively balance a review of basic concepts (biases, such as confirmation bias and detection bias) with information that may be new and interesting to readers at any stage of their statistical education (terminology such as ‘apophenia’, seeing patterns or connections in random data and ‘pareidolia’, when random stimuli are perceived as significant, such as in when one sees the Virgin Mary in a grilled cheese sandwich).

The next chapter also has a highly cautionary tone. The authors discuss how the phenomenon of regression toward the mean can lead to invalid reasoning and offer a quote by John von Neumann to caution the reader on overfitting observations when developing models: “With four parameters, I can fit an elephant, and with five, I can make him wiggle his trunk”. Here the authors are making the point that a good fit to observations is not always the best model, as a model should be flexible in order to capture both systematic and unsystematic patterns. The authors refer to a regression analysis used during World War II to predict the accuracy of bombing; this example clearly demonstrates the effects of misinterpreting regression weights and the variability of a model depending on what variables are included.

Chapter 7 begins with a refresher on populations, samples, distributions, and the central limit theorem, and then progresses to a discussion on the “beauty of natural variation” and the importance of having an “appreciation of random processes”. The authors posit that most people tend to underestimate the amount of variation that should be present in random data. For example, in the context of sports, it is tempting to try to explain randomness in performance as “pressure” felt by athletes, players being “in a slump”, or a team “having chemistry”. One important message in this chapter is that often randomness is misunderstood, and bad things can happen when you don’t see randomness where you probably should (such as in the Bernie Madoff case, where investments consistently and unrealistically gave 12% returns without any variability) or when cause is attributed when none exists (effectiveness of medical treatment in clinical trials). Another important point made in this chapter is in the section about the pitfalls of software, stating specifically that software can provide lots of output, but that does not necessarily mean the output should be used. This point is useful for statistics students to be aware of when interpreting output. The authors state their view that open-source software is preferable over closed-source software packages, since closed-source software allows analyses to be conducted without an understanding of what is really happening. However, in reality, using open-source software may not be a reasonable expectation for all readers, especially for statistics students.

The last chapter of Part I is dedicated to the field of psychometrics. The authors’ discussion of reliability and validity is clear and is supported by simple examples. The example used to explain Cronbach’s alpha relates to the question of whether or not

criminal behavior is a central component of psychopathy. It is an interesting example and very effective for enhancing the readers' understanding of this coefficient.

Part II, on data presentation and interpretation, begins with a short chapter that emphasizes the importance of presentation to help uncover the story behind the data. A list of fourteen common mistakes that an analyst should be aware of and avoid when presenting data is given here. Subsequently, in Chapter 10, the authors address two frequent offenses: underreporting data and misreporting data. The authors take up a quote by Rudy Giuliani that reveals a common confusion between mortality and survival, which leads the authors to explain both of these concepts along with relative risk and absolute risk, and to revisit the importance of framing and providing context when presenting data. Finally, the authors advise readers to know the population surveyed and in particular understand who may have been uncounted in that population.

Chapter 11 is a surprisingly fun chapter discussing internal validity and the Bradford-Hill criteria for establishing causality, with the inclusion of an amusing story about R.A. Fisher. Apparently Fisher disagreed with Hill regarding a link between smoking and cancer. To mock Hill, Fisher wrote an elaborate proof linking apple importation with divorce rates. Next, the authors provide a good explanation of standardization along with a caution to "look under the hood" when making conclusions based on aggregated data. The topic of Chapter 13 is meta-analysis, a type of analysis that has become popular in recent years. The authors provide a thorough explanation of what it is, why it became popular, and of course offer some criticisms of the technique. The authors give supporting examples of problems that arise in the interpretation of meta-analyses and warn against unethically motivated meta-analyses. Wrapping up Part II, Chapter 14 effectively walks the reader through a Supreme Court death penalty case to present the troubling topic of "statistical sleuthing".

Part III explores experimental design and data-collection topics, including general background on types of experimental studies, ethical considerations, and the Federal Rules of Evidence (FRE). Chapter 15 provides an excellent discussion of clinical trials, and of course provides specific cautions a researcher should ask him- or herself when dealing with observational studies. The next chapter is a great reference for statistics and/or public health students, as it walks the reader through several major historical milestones in the development of ethical guidelines for human experimentation. The authors do an excellent job here and throughout the text to incorporate current events into their discussions. In this chapter, to which historical events (Nuremberg Code, National Research Act, Declaration of Helsinki, etc.) lend themselves heavily, they successfully bring more contemporary events into the conversation, such as the apology that former President Clinton gave while he was in office for the circumstances surrounding the Tuskegee Syphilis Study, which makes the text relevant and more interesting than just a history lesson and review. Finally, the section ends with a chapter devoted to issues of admissibility of evidence and expert testimony in court cases. Also an entertaining section, the authors discuss "junk science" and recent examples from news sources to which all readers will be able to relate. The discussion of the Freedom of Information Act (1998) and the Data Quality Act (2001) along with their impacts on the data environment are thought-provoking, helping the reader to recognize recent changes to the environment due to the legislation and to

anticipate changes that will continue as more and more data are made available and new ways to look at these data are realized.

Throughout the text, the authors have provided a plethora of additional resources, with excerpts from court cases, appropriate quotes at the beginning of each chapter, and extensive notes at the end of each chapter. I recommend reading the notes, as some notes are unexpectedly quite humorous and some even contain jokes.

Overall, this book stands out as a unique text that combines a review of mathematical theorems and formulas, guidance on how to be sharp when using or interpreting statistics, and the impacts (often negative) that can happen (and have happened) when analyses are conducted carelessly. All of these components come together effectively to raise the reader's awareness of ethics in creating and interpreting statistics, and ultimately help the reader become a more astute and ethically responsible analyst.