A Reading Guide to Intuitive Biostatistics

Nathan Brouwer

2018-08-13

Contents

Preface

This is a reading guide to Harvey Motulsky's Intuitive Biostatistics: A Nonmathematical Guide to Statistical Thinking, 4th edition. More information about the book can be found at the book's website, http://www.intuitivebiostatistics.com/, and it can be purchased from Amazon.com. Motulsky is the CEO and Founder of GraphPad, a user-friendly statistical software popular in some branches of the life sciences.

Intutitive Biostatistics is a fabulous book for researchers that need to understand or do basic statistics and either need a concise primer on the key issues and/or are turned off by the equations underlying the statistical methods. Instead of using math to explain statistical methods, Motulsky focuses on written explanations, real-world examples, and novel graphing approaches. An excellent aspect of this book is that it unpacks common misunderstandings that researchers have, such as how to interpret p-values (Chapter 17), and signposts bad practices that must be avoided (like p-hacking). Again, this is done by focusing on intuition, not math. Motulsky also presents best practices in plotting, data presentation, and data reporting, emphasizing the key aspect of adequate and accurate presentation of results.

This reading guide serves several purposes:

- Highlight the parts of the book I focus on in my teaching (and so will be on any tests!)
- Provide additional complementary examples
- Indicate extensions or alternatives
- Provide citations and links to resources for follow-up
- Indicate where others (including myself, though I am not a trained statistician) might disagree with Motulsky

Each part of the reading guide is essentially an outline of each chapter with commentary as needed. In some cases I have written a brief initial commentary to put the chapter in context. I will often indicate the Excel or R functions related to methods or calculations; for a fuller treatment see my other guide *An R Companion to Motulsky's* Intuitive Biostatistics. At the end of each chapter are typically references, a list of R and Excel functions needed to carry out the analyses in the book, and study questions to consider.

My most important notes and comments are generally in **bold** or bulleted. When I've riffed on an idea and its not necessarily key I've usually put in in a block quote, like the one below:

> For example, sometimes I've written about a section, and my text is almost as long as the original section!

This is a work in progress and many sections are not yet annotated; feel free to contact me with suggestions or corrections.

Nathan Brouwer brouwern@gmail.com

Chapter 1

"Statistics & Probablity Are Not Intuitive"

Commentary

In this introductory chapter Motulsky sketches out some major reasons why people struggle with statistics and probability. This chapter assumpes some basic familiarity with statistical ideas. Sometimes this chapter is a bit terse - its meant to highlight key ideas, not fully discuss or demonstate them.

Vocabulary

1.0.1   Motulsky vocab

- sample
- population
- Bayesian
- multiple comparisons
- regression to the mean

1.0.2   Aditional vocab

- Bayes theorem
- pre-registration
- exploratory analyses

1.0.3   Key functions

None

Chapter Notes

1.0.4   We Tend to Jump to Conclusions

Motulsky uses the phrase "**generalize from a sample to a population**" without defining
what this means. In general, this means to look at some subset of the world - either something
experienced in real life or generated using a scientific study - and conclude that what was seen in
the subset occurs elsewhere. In the example he uses, his daughter experienced meeting doctors,
and they all were male, so she generalized to the rest of the world that all doctors must be
male. While this example is trivial, anytime we generalize from sample to population (or from
a part to the whole) we run the risk that our sample is biased. It could be biased becauase
we didn't take a good sample, such as relying just on personal experiencce. Or it could be a
rigorously collected scientific sample, but still be non-representative. What if he wanted to prove
his daughter wrong and so randomly selected 10 doctor's offices for a web search and looked up
who the senior physician is. If he happend to find my doctor's office, he'd see that its a women,
Dr. Cathy Lamb. However, it is possible that he could look up 10 doctor's and they could all
still be male.

1.0.5   We Tend to Be Overconfident

1.0.6   We see Patterns in Random Data

1.0.7   We don't realize that coincidences are common

He doesn't use the specific term, but he is alluding to the concept of **hindsight bias**.

1.0.8   We don't expect variability to depend on sample size

Motulsky cites a paper by Andrew Gelman here, one of the most thought provoking - though sometimes
just provoking – statistics bloggers of the last decade. He blogs regularly at Statistical Modeling, Causal
Inference, and Social Science and writes non-technical pieces for a number of outlets, including Slate. He is
also prominent Bayesian.

1.0.9   We Have Incorrect Intuitive Feelings About Probability

1.0.10   We Find it Hard to Combine Probabilities

1.0.11   (We Avoid Thinking About Ambiguous Situations)

(This section appear in previous versions; I am not sure where/if it occurs in the 4th edition)

1.0.12   We Don't Do Bayesian Calculations Intuitively

Motulsky doesn't define **Bayesian** here, though its not central to what he's talking about. In
this example, "Bayesian calculations" refers to a particular type of probability calculation using
Bayes Rule. His example is a classic example of how probability calculations are used for
diagnostic testing.

More generally, "Bayesian"" refers to a particular way to use the mathematics of probability to make inference. All mathematicians agree on the basic rules of probability calculations. In contrast, when it comes to using the math of probablity to make inference from a sample to a population - that is, to do statistics - there is a huge rift between **Frequentists** and **Bayesians**.

1.0.13 We are Fooled By Multiple Comparisons

The study on astroglogical signs here is a great paper intended to "To illustrate how multiple hypotheses testing can produce associations with no clinical plausibility" (Austin et al 2006, Abstract). "Multiple hypotheses testing" means the same thing as "multiple comparisons." As Motulsky indicates, if you test multiple hypotheses or make multiple comparisons between things, sooner or later you'll find a strong association. This is why it important to make specific hypotheses prior to the beginning of a study - ideally even publically **pre-registering** them - and properly indicate which analyses were defined in advance and which are **exploratory analyses**.

Multiple Comparisons is a big topic that Motulsky doesn't go into detail yet. He devotes several excellent chapters to this topic elsewhere. This issue of multiple comparisons is a big and controversial one. For a discussion of multiple comparisons

1.1 We tend to ignore alternative explanations

1.2 We are fooled by regression to the mean

Regression to the mean is a concept that isn't typically taught in intro stats courses, especially for ecology. For its relevance to ecology and evolution see the paper by Kelly and Price (2006) "Correcting for Regression to the Mean in Behavior and Ecology" in *American Naturalist.*

1.3 We let our biases determine how we interpret data

1.4 We crave certainty, but statistics offers probability

1.5 Further reading

1.5.1 References

Austin, Mamdani, Juurlink and Hux 2006. Testing multiple statistical hypotheses resulted in spurious associations: a study of astrological signs and health. Journal of Clinical Epidemiology 59:964–969 Open Access

1.5.2 Annotated Bibliography

1.5.2.1 Multiple comparisons

Bender & Lange 2001. Adjusting for multiple testing—when and how? Journal of Clinical Epidemiology. 54:343–349. Abstract

Multiple comparisons is a thorny issue that Motulsky briefly introduces here in Chapter 1 and discusses in depth elsewhere. Throughout the book Motulsky focuses on the need for multiple comparisons procedures in general, and the most popular ones used; he doesn't go into the broader arguements about their use and the many ways they can be problematic. Bender & Lange (2001) give a taste of the mess made by multiple comparisons issues. They note "...there seems to be a lack of knowledge about statistical procedures for multiple testing. For instance, multiple test adjustments have been equated with the Bonferroni procedure, which is the simplest, but frequently also an inefficient method ..." (pg. 343). They discuss the various positions that have been taken for and against multiple comparisons in the biomedical sciences, and advance their particular perspective on the issue. Elsewhere in the book Motulsky discusses the Bonferonni correction under the heading "The Traditional Approach to Correcting For Multiple Comparisons." He then outlines a more contemporary approach, the **False Discovery Rate (FDR)**. Bender & Lange (2001) was written before the FDR became popular and instead briefly disucuss other alternatives, including Holm modificaiton to the Bonferroni procedures and advanced computational methods.