

Intuitive Biostatistics: A Nonmathematical Guide to Statistical Thinking (rev. 2nd ed.)

Harvey MOTULSKY. New York, NY: Oxford University Press, 2010. ISBN: 978-0-19-973006-3. xxv+447 pp. \$54.95 (P).

A growing number of people need to be able to analyze data and reason quantitatively – in short, to understand statistics. A significant fraction of those people, however, are held back from doing so by a barrier of mathematics. A number of textbooks have made worthy efforts to reach this audience by attempting to explain statistical ideas with a minimal use of equations. Of particular note is the groundbreaking *Statistics* (Freedman et al., 1978), which states: “Mathematical notation only seems to confuse things for many people, so this book relies on words, charts, and tables; there are hardly any x’s or y’s.”

Harvey Motulsky’s *Intuitive Biostatistics* shares that sentiment: “If you like to think with equations and prefer to learn by following mathematical logic, you have plenty of statistics texts to choose from. This book is for the many students and scientists who find math confusing and prefer verbal explanations.” Motulsky’s book is intended to serve as a guidebook for consumers of statistics: scientists, professionals, and graduate students in medical and related fields who need to understand the statistical portions of the journals they read, or who analyze data using software packages and need to be able to interpret the results.

Certainly, there is a need for such books, especially ones that cover important biostatistical topics like survival analysis and case control vs. cohort designs. However, the value of a book that explains statistics verbally depends entirely on how good those verbal explanations are. It is not easy to get statistical ideas across without equations and mathematics, and all of us in the profession have no doubt seen plenty of awful attempts to do so.

So, how good are Motulsky’s explanations? In my opinion, very good. Particularly excellent are his delightful first chapter (available at www.intuitivebiostatistics.com), which contains examples of the many ways in which people’s natural intuition leads them to incorrect statistical conclusions; his series of chapters on p -values, easily the most widely misunderstood and misused statistical concept; and his chapter on the hidden dangers of multiple comparisons. The verbal explanations are nicely supplemented with figures and examples involving real data. The book legitimately emphasizes confidence intervals over hypothesis tests, instead of the lip service one often encounters. Special attention (and

clear typesetting) is given to key assumptions that may be violated and that should be considered in the application of specific methods. Furthermore, these explanations get right to the point and make that point very concisely, allowing the slim 447-page paperback to cover a tremendous amount of ground. In addition to all the topics you'd expect to find, one also encounters, for example, nonlinear regression, false discovery rates, Bayesian logic, noninferiority testing, and bootstrapping.

Nevertheless, it must be pointed out that the juxtaposition of Motulsky's and Freedman et al.'s books at the beginning of this article is misleading – they are quite different books written by quite different authors for quite different purposes. Dr. Motulsky is not a statistician by training (his Ph.D. is in pharmacology), and his approach will no doubt strike many statisticians as strange. For example, the terms “distribution” and “sampling distribution” are absent from the index. Probability is never defined, and only very briefly touched upon. While his understanding and explication of the big picture, and of the practical aspects and pitfalls of biostatistical methods are quite good, the book falters occasionally when wandering into more theoretical and mathematical territory, such as the frequentist interpretation of confidence intervals. After mentioning that it is incorrect to say that there is a 95% chance the population value lies within the CI, it goes on to say, “Instead, it is proper to say the following: ‘There is a 95% chance that this CI contains the population value’. This is a fine distinction that is unlikely to affect your understanding of statistical results.”

The preceding paragraph is likely to draw the ire of many a statistician, but I would urge the reader not to be so quick to dismiss the book. Motulsky may not be a Ph.D. statistician, but as the author of GraphPad (a statistical analysis and graphics package) and a textbook on nonlinear regression (Motulsky and Christopoulos, 2004), he is also no dabbler in the field. This book arose from his efforts teaching statistics to first-year medical students at the University of San Diego, and between that experience and his hours spent fielding GraphPad user questions, he has seen a tremendous variety of statistical questions and misconceptions. These are on display in his “Q & A” sections, which range from the subtle “Q: Is the P value the probability that the null hypothesis is true? A: No. The P value is computed assuming that the null hypothesis is true, so it cannot be the probability that it is true.” to the much less subtle “Q: Can P values be negative? A: No. P values are fractions, so they are always between 0 and 1.” It would not have occurred to me that people might wonder whether p -values can be negative, and to go out of my way to explain that. Like most statisticians, my natural instinct would be to define p -values with an

equation and just assume it was obvious to everyone that it couldn't be negative. But clearly, it isn't obvious to everyone (first-year medical students, I'm looking at you).

I have spent several semesters teaching introductory biostatistics to an audience of graduate students in public health, physicians, nurses, pharmacists, toxicologists, and other health professionals. In past semesters, I used *Fundamentals of Biostatistics* (Rosner, 2006), the kind of introductory Biostatistics text written by and appealing to statisticians, who, as Motulsky put it, like to think with equations and follow mathematical logic. My students hated it. Many told me that they had stopped reading it altogether, and were relying instead on books with titles like "Statistics for Dummies."

I didn't want my students reading books for dummies; despite struggling with math, they were thoughtful people, and I wanted my students reading a thoughtful book. So last semester, I suggested *Intuitive Biostatistics* to them as a supplementary text, and the feedback I received was quite positive (though not a random sample, so take this for what it's worth). I intend to require it next semester, and would suggest that others who teach such a course consider the idea. For myself and I would imagine virtually all statisticians, the mathematical logic is the easy part of teaching introductory statistics. Coming up with clever analogies, examples, and verbal explanations that make sense to non-mathematical students is much harder. *Intuitive Biostatistics* is filled with these sorts of explanations. Its author understands statistics, but it doesn't sound like it was written by a statistician, and that's what makes the book so effective as a supplementary text. Students can hear two explanations: one from a statistician, and one from someone who speaks their language. That seems to be exactly what many people need.

Patrick BREHENY
University of Kentucky

REFERENCES

- Freedman, D., Pisani R., and Purves, R. (1978), *Statistics*, New York: Norton.
- Motulsky, H. and Christopoulos, A. (2004), *Fitting models to biological data using linear and nonlinear regression*, New York: Oxford University Press.
- Rosner, B. (2006), *Fundamentals of Biostatistics*, Belmont, CA: Duxbury.