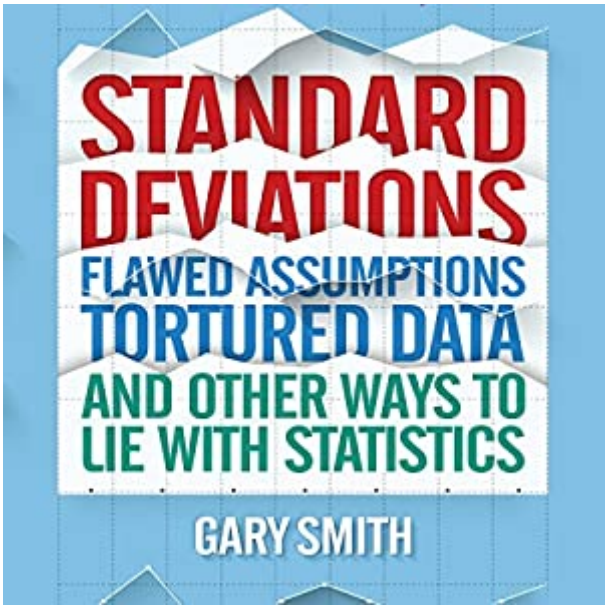
Standard Deviations: Flawed Assumptions, Tortured Data, And Other Ways To Lie With Statistics



Synopsis

Did you know that baseball players whose names begin with the letter "D" are more likely to die young? Or that Asian Americans are most susceptible to heart attacks on the fourth day of the month? Or that drinking a full pot of coffee every morning will add years to your life, but one cup a day increases the risk of pancreatic cancer? All of these "facts" have been argued with a straight face by credentialed researchers and backed up with reams of data and convincing statistics. As Nobel Prize-winning economist Ronald Coase once cynically observed, "If you torture data long enough, it will confess." Lying with statistics is a time-honored con. In Standard Deviations, economics professor Gary Smith walks us through the various tricks and traps that people use to back up their own crackpot theories. Sometimes, the unscrupulous deliberately try to mislead us. Other times, the well-intentioned are blissfully unaware of the mischief they are committing. Today, data is so plentiful that researchers spend precious little time distinguishing between good, meaningful indicators and total rubbish. Not only do others use data to fool us, we fool ourselves. With the breakout success of Nate Silver’s The Signal and the Noise, the once humdrum subject of statistics has never been hotter. Drawing on breakthrough research in behavioral economics by luminaries like Daniel Kahneman and Dan Ariely and taking to task some of the conclusions of Freakonomics author Steven D. Levitt, Standard Deviations demystifies the science behind statistics and makes it easy to spot the fraud all around.

Book Information

Audible Audio Edition
Listening Length: 9 hoursÃ Â andÃ Â 19 minutes
Program Type: Audiobook
Version: Unabridged
Publisher: Gildan Media, LLC
Audible.com Release Date: June 1, 2016
Whispersync for Voice: Ready
Language: English
ASIN: B01G61WNTW
Best Sellers Rank: #32 inÃ Â Books > Audible Audiobooks > Science > Mathematics #484 inÃ Â Books > Science & Math > Mathematics > Applied > Statistics #828 inÃ Â Books > Science & Math > Mathematics > Applied > Probability & Statistics

Customer Reviews

Like the previous reviewer, I have a Ph.D. in a related field (finance) and also like him, I had never heard of the author, Dr. Smith, before seeing this book last week. After that we are alike in no apparent way. I heard no NPR interview, but when I saw this book on the shelf, I flipped through it and liked its contents well enough to buy it. I did not research the author's CV to decide whether his logic is sound; when a publication is in my area of expertise, I prefer to make that evaluation myself. I am only about a third of the way through the book, but so far I sufficiently impressed that I will probably make it assigned reading the next time I teach a doctoral class that focuses on methodological issues. If you're not doctoral student, don't worry; the book is written for a reasonably bright layman, but many of the issues Dr. Smith raises are common (but not obvious) errors I have seen many researchers make, and I really like seeing so many of them summarized in one place.Indeed, I find it difficult to understand the frame of reference from which the previous reviewer sees the book. He starts by saying "most of the stuff Smith covers has been extensively documented elsewhere," but then goes on to imply Smith is unqualified to summarize said documentation. He criticizes Dr. Smith because "it [sic] like he enjoys taking pot shots at other more respected economists," but then himself asserts that "[Dr. Smith's] CV implies that he was denied tenure at Yale," and then refers to his "bitterness" and his "being willfully blind to the real world." "Enjoys taking pot shots?" Physician, heal thyself. Let me address just the implications of the conjecture that Dr. Smith was denied tenure at Yale. Yale is a world-class university, and Dr. Smith was undoubtedly selected from a set of a hundred or more applicants from other top universities. The fact that Dr. Smith was selected from that pool strongly signals to me that he is exceptionally well trained and almost certainly sufficiently well-qualified to write such a book.I do thank the previous reviewer for pointing me in the direction of "Murmane [sic] and Willett's" book, which he suggests summarizes "plenty of rigorous statistical methods for making valid inferences from observational data." This is a big issue in finance and economics [e.g., unlike a chemist who can control temperature or a physicist who can control velocity, a finance professor cannot tell Apple to fire its CFO so we can see what the effect on share price is], and the most common way of trying to deal with it is with instrumental variables (IVs). However, while this approach is theoretically quite sound when the right IV is available, in practice it is typically difficult to find an appropriate IV. My own personal suspicion is that using an imperfect IV creates at least as many problems as it solves (indeed this is a research area I intend to pursue myself when time permits), but I look forward to reading Murnane and Willett's analysis of this issue.OK, enough about the previous review, what about the book itself? The author is quite correct that, given the pressure to publish, many researchers will be inclined to look at different combinations to try to find statistically significant

relations between variables of interest, and that this changes the statistical significance level when such a relation is found. We all see the results of this regularly. Last year eggs were bad for your health; this year they are good. Last month vitamin pills were good; this month they are bad. The traditional wisdom is that because protein is hard on the kidneys, feline renal problems should be treated with a low-protein diet, but there is a growing body of research (e.g., see Hodgkins' book Your Cat, a real bargain at the Kindle price of $8.89) suggesting that in the wild, a cat's diet is about 2/3 fat and 1/3 protein, and when we give them the low-protein diet characteristic of dry food we are actually causing renal (and other) problems, not solving them. Is coffee good or bad? It depends on whether you are reading today's newspaper, or yesterday's, or last week's. The list of such flip-flops is endless, and indeed the medical journals seem to be taking the lead in addressing this problem (e.g., Ioannidis' Why Most Published Findings are False, http://www.plosmedicine.org/article/info%3Adoi%2F10.1371%2Fjournal.pmed.0020124).Why are all these tests so inconsistent? Part of the reason is that many are based on observational data and, the IV method notwithstanding, it is difficult to determine cause and effect from observational data. Part of the reason is that sometimes researchers will keep looking until they find a statistically significant result that supports their hypothesis, and then use the same data to "confirm" that hypothesis. Sometimes there was a bias in the way the samples were selected. Sometimes the sample size is too small, or the data were grouped incorrectly, or the wrong data were excluded. If you want a better understanding of these issues, and to be better prepared to assess for yourself what you see in the media, I highly recommend this book.

This makes statistics fun & you learn so much, especially how you are being fooled everyday. I bought 3 more copies & sent them to friends, some of them MBA's & big-time stock buyers & they too have learned a lot.

I started to give this book 2 stars, but couldn't honestly say I didn't like it.It is a very enjoyable read - until it isn't, and for me that occurred about ÃƒÂ¢Ã â |Ã â • of the way through. At that point the malarkey factor made me ask myself why I was taking time to read i. My answer was that Smith is an engaging writer, and I kept oping that he would explain some of the tings he was point at derisively. Do athletes hat hot streaks? Sure, and the good ones manage to have them occur during Olympic competition.Why spend a number of pages considering whether a person with initials GOD has any predictable charactereristic different that a person with initials PIG. Do we care about Houdini or psychic phenomena?There is a lot of good material here, written in an engaging way.

The bottom line is to b skeptical of theories that are "proven" by the same data that was used to find them. Point well taken, and very useful.But there are somereal problems. The author repeatedly decries any chart that does not show the zero point on the ordinal axis. But if you want to look at a large number, such as CDP, taking the axis down to zero yields a meaningless chart, mostly empty but with a straight line at the top. A meaningful presentation requires that the data curve occupy a sizable amount of the graph.There is repeated reference to the independence of individual trials, and a phenomenon he call "Regression to the Mean". He suggests that if students take a test, you might see one do poorly, and another do very well.The he asserts that on a retake of the test the poor studied will have a higher score, and the good one a lower score as both regress toward the mean.I suspect the Professor Smith is an engaging teacher. I also suspect that he uses material like this to relieve the stress of discussing degrees of freedom, and the math of a hyperkurtoid normal distribution.It is a good read, but leaves the reader with a degree of suspicion about any paper describing empirical results. I had always assumed that these problems were within the purview of editors, and peer reviewers prior to publication, and independent replication of results after publication, but before general acceptance by the relevant community.And of course any change in prevailing models is subject to the resistance of a paradigm shift (as described by Thomas Kuhn).Caveat lector.

A lot of great examples of how data gets used incorrectly. Nothing super advanced, but really valuable for someone starting off in Statistics

Excellent insights on understanding data analysis.

Very enjoyable and enlightening

This is a good book about how one might get fooled by manipulative presentation of data.

A number of books have been written over the past decade or two on the abuse of statistics. The reasons for such abuse seem to vary from simple ignorance of proper statistical techniques to deliberately mislead in order to fulfill some hidden agenda. This book is a further addition to this growing and useful library.In this case, the author, a professor of economics, covers much of the same material as prior authors collectively have; that is, this book covers a lot of ground ÃƒÂ¢Ã â ¬Ã â œ more than individual prior authors that I am familiar with. Much space is

particularly devoted to the stock market.I found the authorÃƒÂ¢Ã â ¬Ã â„¢s prose to be really quite captivating. It is also lively, often witty, accessible and authoritative. Although most of the discussions were, in my view, engaging and understandable, in a few cases I either could not see the authorÃƒÂ¢Ã â ¬Ã â„¢s point or simply disagreed with his conclusion based on the information presented. Also, very occasionally, I found that explanations given here on certain specific topics were not as clear as those given on the same topics by prior authors whose books I had read. But despite these few relatively minor shortcomings, I did thoroughly enjoy this book, as should anyone interested in how statistics can be misused.

(Predictive Analytics, Data Visualization, Data Analytics, Business Analytics, Decision Analysis, Big Data, Statistical Analysis) Discovering Knowledge in Data: An Introduction to Data Mining (Wiley Series on Methods and Applications in Data Mining)