The Statistics Teacher Network

Number 48 | ASA/NCTM Joint Committee on the Curriculum in Statistics and Probability | Spring 1998

Book Review...

Statistics

Third edition

David Freedman, Robert Pisani, and Roger Purves, University of California, Berkeley

1998, 578 + A-123 pages, ISBN 0-393-97083-3 New York: Norton

http://www.wwnorton.com

This book was a pioneer when it first charted a new course for introductory statistics two decades ago, and it remains a leader today. All teachers of statistics, at all levels, can learn from it and should read it. If you want a textbook for a course whose content matches what this book covers, you can't do better. No other book even comes close.

The first edition of this book came out in 1978, at a time when computing meant batch jobs on a mainframe. This was eight years before the current NCTM standards, eleven years before Moore and McCabe (1989), and light years before the first AP exam in statistics. In the time since then, change in statistics education has been so fast and has cut so deep that it can be hard to appreciate how far ahead of its time this book was when it first came out. What made the book so revolutionary? Three things: data, exercises, and exposition.

Real data

These days real data in a textbook has become the norm, although you can still tell the better books from the weaker ones by what they actually do with the data. (Weaker books use data only to illustrate the mechanics of the methods; better books also reverse priorities, using examples to show how the methods help you understand your data.) Two decades ago few books included much real data. There were no desktop computers or graphical interfaces; even hand-held calculators were very new, quite expensive, and extremely limited in what they could do. One way to appreciate the data revolution in statistics teaching that has so changed things in the last decade is to turn the clock back on exploratory data analysis (EDA), to the time when transforming to logs meant using a "break table" (Tukey, 1977, pp. 62, 71-72). Then, as now, good data analysis often requires transforming to a new scale, to ensure that the data conform to the assumptions required by whatever methods you are using. In the "early days" — as recently as the late 1970s when *Statistics* appeared — transforming meant going through a painfully slow and tedious process of looking up each data value in a table to find its logarithm, root, or reciprocal. No wonder hardly anyone taught data transformations! Tukey's introduction of break tables began to change that, by making it possible to carry out transformations "by hand" much more quickly. These days, when logs and roots are just a key-press away and the technology to display box plots is smaller than a sandwich and cheaper than a pair of Nikes, we tend to think of EDA with the emphasis on "exploratory." In the old days, however, one of the effects on statistics education was to put new emphasis on "data" as well. Tukey's methods made it possible to do serious analysis of real data without a computer. For statistics

Also In This Issue...

education, this was the analog of translating the Latin mass into the vernacular—suddenly anyone could work with real data, not just the high priests of the mainframe.

EDA has largely gotten the recognition it deserves, but there was also a less-recognized parallel development, not in statistics generally, but in statistics education: Freedman, Pisani and Purves wrote an introductory book that showed how to make systematic use of real data in examples and exercises that dealt with ideas rather than formulas and arithmetic. Here, too, the effect was to make it possible for beginning students, without the aid of technology, to wrestle with the angel of real data. The innovation was not exclusively due to Freedman et al., but no previous book had made such systematic and effective use of it. The easiest and most direct way to see this in action is to look at their exercises.

Exercises

These remain distinctive 20 years after the first edition appeared. Here's a (non-random) sample.

> p. 263, #1. In the US in 1990, 20,273 people were murdered, compared to 16,848 in 1970 — nearly a 20% increase. "These figures show that the US became a more violent society over the period 1970-1990." True or false, and explain briefly.
>
> p. 429, #5. Before the strike of 1994, the median salary of the 746 major league baseball players was about $500,000. The lowest salary was about $100,000 and the highest was over $5,000,000. Choose one option and explain:
>
> (i) The owners were paying out around 746 x $500,000 = $373 million per year in salaries to the players.
>
> (ii) The owners were paying out substantially less than $373 million per year to the players.
>
> (iii) The owners were paying out substantially more than $373 million per year to the players.
>
> p. 430, #10. Pearson and Lee obtained the following results in a study of about 1000 families:
>
> average height of husband 68 inches, SD 2.7 inches, average height of wife 63 inches, SD 2.5 inches, r .50.
>
> Among men who were about 5 feet 4 inches tall, estimate the percentage who were shorter than their wives.

All these exercises are based on real data. None requires much arithmetic. Each requires substantive engagement with important elements of statistical thinking. A student who works through a book full of exercises like these will not waste much time grinding away in the Mills of Algorithm, but will instead learn a lot about the logic of statistics. Judge a book by its exercises, and you cannot go far wrong.

Exposition

The third distinctive feature of the book by Freedman, Pisani, and Purves is its exquisitely crafted exposition, in non-technical language, of difficult mathematical ideas, especially those that support the bridge from probability to inference. Of course there are many books that cover essentially the same mathematical content, but at a higher technical level. There are also other superb books written in non-technical language, such as, at a lower level, David Moore's (1997) *Statistics: Concepts and Controversies*, and at a higher level, Iversen and Gergen's (1997) *Statistics: The Conceptual Approach*. Sadly, there are also popular books that somehow manage to be full of trees and underbrush without ever suggesting the presence of a forest. For example, consider the question of when to use a z test, when to use a t test, and when to use some other test. Both FPP and another widely used book (Triola, 1995) use flowcharts to tell you what to do. The chart in Freedman et al. (p. 494) has distilled things to their essence: there are no symbols, and a spare total of just 63 words. A good student can look at their flow chart, see the main ideas, and remember them. The flow chart in the other book (p. 446) is so full of symbols that the number of subscripts comes to 63 — that author needs a subscript for every word in the FPP chart! Worse, the advice encoded in Triola's dense page of symbols is wrong: wrong about when to use the z test, wrong about when to use non-parametric methods, and wrong to recommend a preliminary F test to check whether population standard deviations are equal.

Freedman, Pisani and Purves not only get things right, they have thought harder than most authors about how to say things so that students learn to get them right, too. The exposition has been described as "friendly" and "like a detective novel," but the authors don't shy away from substance. Here, for illustration are the contents of two boxed summaries taken from the sections on correlation and regres-

sion. To appreciate more fully what the authors have wrought here, you might test yourself before reading ahead: can you define the correlation, using only 15 words, in a way that students can easily remember and can also use as a formula for calculating the value of r? Can you then use another two short sentences to tell how r characterizes the quantitative relationship between two variables? Here's how Freedman & Company do it:

> Convert each variable to standard units. The average of the products gives the correlation coefficient. (p.132)

> Associated with each increase of one SD in x there is an increase of only r SDs in y, on the average. (p. 160)

In just three sentences they have given a definition-cum-computing-rule, and told how the correlation is related to the fitted regression slope (although the latter is not named in this summary.) Here are two more boxed summaries related to correlation and association:

The correlation coefficient is a pure number, without units. It is not affected by interchanging the two variables, adding the same number to all the values of one variable, multiplying all the values of one variable by the same positive number. (p. 143)

> If a study draws conclusions about the effects of age, find out whether the data are cross-sectional or longitudinal. (p. 60)

The rest of the book is written to the same high standards.

Topics covered

This book is truly magnificent in its presentation of those ideas from the Advance Placement syllabus that deal with probability and inference, precisely the topics that are hardest for students to understand. At the same time, however, the authors pay little or no attention to a number of topics listed in the Acorn book (College Board, 1996). There are no dot plots, stem plots, or box plots. The interquartile range gets only two sentences (p. 89): "By definition, the interquartile range equals '75th percentile - 25th percentile.' This is sometimes used as a measure of spread, when the SD would be too heavily influenced by a small percentage of cases in the tail, of the distribution." Residual plots are described carefully and illustrated briefly, but a page is about all you get; influential points aren't mentioned. There's also nothing about transforming to logs, roots, or reciprocals.

All these omitted topics are part of the Advanced Placement syllabus, and all are important. Regarded as a handful of mere techniques, they are almost all easy to teach, and as mere add-ons, they are also easy to dump into a book (or into a course). The authors have thoughtfully resisted the temptation, knowing that although it is easy to pour a little cold wine over a meal after you've cooked it, that doesn't make it French cuisine. Taught just as tools, box plots and the like might increase sales of the book, but only to those who can't tell surface from substance. Tools like these belong in a book only to the extent that the authors devotes serious attention to exploration — pattern searching without an underlying probability model — and diagnostics — empirical checks of whether the data and a proposed model are at odds with one another. Freedman et al. have written a book that marches to a different drummer. In their book the time that students might have spent on exploration and diagnostics is committed instead to achieving a better understanding of the logic of inference. The priorities reflected in the AP syllabus are somewhat different: following that program gives more experience with empirical aspects of data analysis, but the inevitable tradeoff means a less sure grasp of the logic of inference.

What's new in this third edition? The changes can be quickly summarized using the same three categories that characterize the first edition's innovations: data, exercises, and exposition. Data sets have been brought up to date. The exercises have been expanded not only by sprinkling new ones throughout the book (about 200, according to the preface), but also by creating four sets of comprehensive review exercises that cover not just the preceding chapter, but all chapters prior to that point in the book. Finally, exposition is largely unchanged, both as to style and organization, although there are new sections on linear change of scale, and on deciding whether dependence makes the two-sample t-test inappropriate.

Conclusion

A lot has changed since 1978. Elementary statistics education can now claim an entire constellation of bright stars that together mark out a convex hull of excellence. For curriculum, we have David Moore's books (the two men-

tioned previously, plus *The Basic Practice of Statistics*, 1995), Siegel and Morgan's *Statistics and Data Analysis: An Introduction*, the book by Iversen and Gergen (1997), and Robert Wardrop's (1995) *Statistics: Learning in the Presence of Variation*. For pedagogy, we now have a great variety of approaches, via Allan Rossman's (1996) *Workshop Statistics: Discovery through Data*, Scheaffer et al. (1996) *Activity Based Statistics*, Spurrier et al. (1995) *Elementary Statistics Laboratory Manual*, Chatterjee et al. (1995) *A Casebook for a First Course in Statistics and Data Analysis*, and Paul Velleman's (1998) *ActivStats*. All these are major contributions, but all are new within the last ten years. In contrast, *Statistics*, by Freedman, Pisani and Purves first appeared 20 years ago. It was a pioneer at the time, and it remains a leader today.

References

George W. Cobb Mount Holyoke College South Hadley, MA 01075 413-538-2401

College Board (1996). *Advanced Placement Course Description: Statistics*. Princeton, NJ: The College Board.

Iversen, Gudmund, and Mary Gergen (1997). *Statistics: The Conceptual Approach*, New York: Springer-Verlag.

Moore, David S. (1995). *The Basic Practice of Statistics*, New York: W. H. Freeman.

Moore, David S., and George P. McCabe (1989), *Introduction to the Practice of Statistics*, New York: W.H. Freeman. 2nd ed. 1993.

Moore, David S. (1989), *Statistics: Concepts and Controversies*, New York: W.H. Freeman. 3rd ed. 1997.

Rossman, Allan J. (1996). *Workshop Statistics: Discovery with Data*, New York: Springer-Verlag.

Scheaffer, Richard L, Mrudulla Gnanadesikan, Ann Watkins, Jeffrey A. Witmer (1996), *Activity Based Statistics*, New York: Springer-Verlag.

Siegel, Andrew F. (1988). *Statistics and Data Analysis: An Introduction*, New York: John Wiley and Sons.

Siegel, Andrew F., and Charles J. Morgan (1996). *Statistics and Data Analysis: An Introduction*, 2nd. ed., New York: John Wiley and Sons.

Spurrier, John, Don Edwards, and Lori Thombs (1995). *Elementary Statistics Laboratory Manual*, Belmont, CA: Wadsworth.

Triola, Mario F. (1995), *Elementary Statistics*, 6th ed., Reading, MA: Addison-Wesley.

Tukey, John W. (1977). *Exploratory Data Analysis*, Reading, MA: Addison Wesley.

Velleman, Paul F. (1998). *ActivStats*, Reading, MA: Addison Wesley Interactive, 1997.

Wardrop, Robert L. (1995). *Statistics: Learning in the Presence of Variation*, Dubuque, IA: William C. Brown, 1995. *Note: Although Brown has dropped all their statistics textbooks, copies of this one are still available from the author.*

Reviewed by George W. Cobb
Mount Holyoke College
South Hadley, Massachusetts
GCobb@MtHolyoke.edu

Book review...

Statistics with the TI-83

Gloria Barrett

1997, ISBN 1-887050-31-11,
Meridian Creative Group
800-530-2355
www.meridiancg.com

We hope to use technology to both enhance and excite the study of statistics. Statistical computations are very cumbersome. Machines do the drudgery of computations in a quick and accurate manner. The graphing calculator offers a flexible portable format to do statistical computations and the TI-83 puts statistical processes at our fingertips. It is easy to understand why there has been a quick marriage between the TI-83 and the study of statistics.

I have been teaching statistics at Hudson High School for the past eight years. The entire period has been exciting. We first used a handheld calculator as our technology, the we got MINITAB, then the TI-81. I thought I had died and gone to heaven when the TI-82 introduced lists and links. I along with many others started writing and linking programs and data sets. The TI-83 was almost a natural extension. I began to realize that statistics could be taught and learned from a data driven conceptual standpoint that was enjoyable. Now we have AP Statistics introduced at a time when the technology is widely accepted. It is an exciting time in statistics education.

All of us only have 24 hours in a day and we do not wish to torture our students so we hope that the emerging technology is easily learned. For the most part, a statistics teacher would be

able to fumble around with the TI-83 and figure out how to incorporate it in many ways into their present courses.

However, good technology helps us use it to generate new knowledge, present ideas in different ways, take thoughts to deeper levels, and save time. Gloria Barrett's *Statistics with the TI-83* helps the statistics teacher and student in that process.

The primary audience for this book would be statistics instructors who are familiar with the basic concepts of statistics and have used the graphing calculator. Instructors who are preparing to teach AP Statistics for the first time would benefit from this book because it gives a relatively thorough course in statistics along with presenting the many capabilities of the TI-83.

I believe that it would also be a valuable resource for any high school or college statistics teacher.

I have personally referenced *Statistics with the TI-83* very often during the year. I first found that the three appendices were very helpful in understanding some of the list upgrades on the TI-83 over the TI-82. I was especially interested in using programs to store lists and naming and storing lists in the SETUP EDITOR. I found this book more informative than the TI-83 manual. The chapters of the book are organized in sequential and topical order to many typical statistics textbooks. Therefore it was easy to reference how to use the TI-83 in teaching and in doing problems similar to the ones that I was currently teaching.

As previously mentioned, the topical format of the book mirrors a typical statistics text. Although each section mentions important topics and ideas related to the section, there is very little attempt to explain the topics. However, many suggestions are given on how to use the TI-83 in exercises to help drive home points of emphasis. I found this very helpful in preparing lessons on various topics. I am always looking for creative ways to teach statistical concepts and browsing the pages has given me numerous ideas.

The pages are very well organized and segmented which is quite helpful for a book that is to be used as a supplemental reference source. The calculator command sequence is very clear and there are usually side notes which help the reader understand what the commands are telling the calculator to do.

I found the side notes that are highlighted on the borders of the pages very helpful. These notes prompt the reader for potential calculator operation problems, conceptual statistical connections to the calculator functions, and references to earlier explanations on how to use a particular function on the calculator. The text is also full of display screens to help the reader see what to expect on the calculator.

Each topic is concluded with a short list of classroom exercises. I was happy to see that many of these exercises did not necessarily teach how to use the calculator, but rather how the calculator and the results obtained from the machine could be used to learn statistics. Many questions were extensions that inspired me to think about the possible answer and then use the TI-83 to verify the results. I have used the classroom excersises as lab type activities in my classes and have been pleased with the results. In some cases gaps in my statistical knowledge were cleared up because of referencing this book. In most cases the points were made clear because the author has no intention of elaborating or going into great detail. Therefore the concept is presented very quickly and clearly in the text with the emphasis placed on using the calculator on related exercises for concept development.

I was suprised that an index was not included at the end of the text nor was a complete command guide for the TI-83. Given that this book seems to be best suited as a reference book or supplemental text, I would have expected both. Since the book is laid out almost identically to many statistics texts, the lack of index has been less of a problem than I originally expected and I realize that the book's table of contents suffices in most instances. I would strongly recommend having the TI-83 users guide available along with this book. Occasionally I found the quick home programs a little confusing, especially the type that included sequences to create simulations. I believe a little more description would be helpful. However, the problems were conceptual rather than practical. I found that the calculator performed quite well once the commands were typed in. The data were produced and the statistical ideas developed. My only confusion was in why the particular command sentence made the calculator do what it did. I noticed that my students were a little more willing to blindly trust that the calculator "knew what to do" than I was.

In summary, I have found that *Statistics with the TI-83* a very helpful supplement to both my statistics text and the TI-83 users guide. I always have all three books around as I plan lessons and activities. I would recommend it as a supplement to the text and guide book and also as a quick workbook for those who are just beginning teaching statistics (especially for AP Statistics) and wish to enhance their knowledge of the subject along with learning ways to use the calculator properly in the teaching and learning of statistics.

Reviewed by David Spohn
Hudson High School
Hudson, Ohio

Book review...

Exploring Linear Relations

Gail F. Burrill
and Patrick Hopfensperger
1998, 208 pages (teacher) 132 pages (student); ISBN 1-57232-211-x (teacher)
ISBN 1-57232-210-1 (student)
Dale Seymour, 800-872-1100

"Why are we learning this? When will we ever use this?" Algebra teachers, have you heard this before? With the module, *Exploring Linear Relations*, you can finally show them why they are learning this and exactly when they will use it. They can see how the relationship among variables and symbols has a basis in the real world.

Exploring Linear Relations is part of the Data Driven Mathematics series. It was written by Pat Hopfensperger, high school mathematics teacher from Wisconsin, and by Gail Burrill, former Wisconsin mathematics teacher and past president of NCTM. The intent of the book is to supplement the algebra curriculum by focusing on the use of real data to motivate the traditional mathematics topic of linearity.

Organization

Exploring Linear Relations is divided into two basic parts. Unit One is the study of linearity. Four lessons introduce the idea of the rate of change, finding and interpreting the meaning of slope, finding the equations of lines, and equivalent lines. Unit Two deals with lines and scatter plots. The students investigate the relationships between two variables by reading and interpreting scatter plots, recognizing different forms of the same equation, using intercepts and rate of change to analyze a situation, finding a "good" line by considering measure of error, drawing median fit lines, determining the strength of a linear relationship by finding a measure of correlation, and understanding the connection between causation and correlation.

Each lesson is designed around a real life problem or mathematical situation. The lesson starts with a series of questions or scenarios that create student discussion and interest in the problem. This is followed by discussion designed to clarify the initial question and "shape" the direction the lesson will take. Students then begin to investigate the situation (individually, in a small group, or with the whole class). Many of the lessons can be divided so that groups of students can do individual parts of the lesson and then put the parts together for final analysis with the class. Often multiple solutions are possible and multiple strategies for arriving at those solutions are encouraged. The role of the teacher is to help students see the "big picture" and to validate and clarify student thinking.

Personal Reactions

As an algebra teacher, I was able to pilot this series before it was published. I found several things in this process. I began my unit on lines with this text and found the students to be somewhat confused. I felt that the confusion was due to the students' unfamiliarity with the idea that a line could have a meaning and was something that had characteristics that gave information. To help to clarify this idea, I went back to my textbook and taught a few of the "traditional" concepts of a line. I then went back to the *Exploring Linear Relations* book and was delighted with the results. My students loved the real life data sets in the book. Rate of change was introduced through the changing cost of cars, gasoline prices, cellular phone usage, and fast food revenues. Other topics include a comparison of median earnings of men and women, comparison of growth rates of babies, sports statistics, TV ratings, food data and a variety of other data sets that held student interest and made them want to work on the activity. I found the marriage of the "traditional" topic taught with a fresh, new approach to be very stimulating to the students (and to me).

One concern I had as an algebra teacher was how I would find the time to teach "something else" in the already packed algebra curriculum. *Exploring Linear Relations* states that the book will take five weeks to complete. I worried about the time I would need to spend, but

found that by carefully selecting the lessons, I could use the time I would ordinarily have spent on lines—plus about an additional week—and catch the flavor of the material, have real student interest, and accomplish the same objectives that are part of my district's course of study. It wasn't an addition of material, but rather a replacement of a chapter in my text with material that evoked students' interest and understanding and still "covered" the necessary content.

Teacher's Guide

The teacher's guide not only gives "answers," but also gives approaches to the topics. Quizzes for selected lessons are included as well as an "end of the module" test. Black line masters that match the activity sheets in the book are included in the teacher's guide.

A scientific calculator is necessary to work through the lessons in this book. Spreadsheet software or graphing calculators with list capabilities are not absolutely necessary but would be extremely helpful. In the Teacher's Guide a section entitled "Procedures for Using the TI-83 Graphing Calculator" is included. There are also instructions for using a spreadsheet for Microsoft Works. Also included are floppy disks written in both the Macintosh and the DOS or Windows format. The data sets from the text are included on these disks.

Conclusion

This module can be used 1) to replace a chapter in the regular algebra textbook; 2) to build understanding of what graphing concepts really mean (to be used after the text material is completed); 3) to be part of a course that includes other Data Driven Modules (specifically to be used following Exploring Symbols and Mathematics in a World of Data); 4) as an introductory unit on slope graphing, and equations in any level of mathematics course; and 5) as part of an integrated mathematics course.

The text deals with students using real life data to solve problems. Hopefully they will learn to think about what the data really mean and how the data are related to the problem. They will have answers to the questions, "Why do we have to learn this? and "When will we ever use this?"

Reviewed by
Maria Mastromatteo
Brown Middle School
Ravenna, Ohio

From the Editor...

As statistics continues to be integrated into mathematics and science curricula, there is increasing need for statistics workshops. ASA would like to help you plan workshops for your district. I am currently working with a district who will be offering four workshops, summer 1999, in Elementary Quantitative Literacy for its K-6 teachers, QL for 7-10, Data Driven Mathematics for 9-12, and Science Education and Quantitative Literacy for its high school science teachers. Their teachers who are interested in AP Stats will attend a workshop at the regional level. Contact me if you are interested in ASA's help in organizing such a program for your district.

Keep NCTM's 1999 annual meeting in San Francisco in mind as there will be a QL CWAC. The day will begin with an address by Dick Scheaffer, followed by breakout sessions in EQL, QL, DDM and AP Stats, and will reconvene for a closing address by Jim Landwehr. It's an exciting program organized by the ASA/NCTM Joint Committee on Curriculum in Probability and Statistics grades K-12.

Have a very productive summer. See you in the fall.

Keep Us Informed...

The Statistics Teacher Network is a newsletter published three times a year by the American Statistical Association—National Council of Teachers of Mathematics Joint Committee on the Curriculum in Statistics and Probability.

We need your letters, announcements, articles, and information about what is happening in statistics education! Please send hard copy, and, if possible, a disk written in standard ASCII text to the editor:

Jerry Moreno, Dept. of Mathematics, John Carroll University, University Heights, OH 44118

or **moreno@jcvaxa.jcu.edu**

or **fax: (216) 397-3033**

Production Editor: *Derek Lawlor, American Statistical Association*

To be added to the mailing list or make an address change, please send your name and address to: *Statistics Teacher Network*, c/o American Statistical Association, 1429 Duke St., Alexandria, VA 22314-3415; (703) 684-1221; fax: (703) 684-2036; E-mail: chris@amstat.org

Printed in the U.S.A

Statistics Teacher Network
c/o American Statistical Association
1429 Duke Street
Alexandria, VA 22314-3415
USA

Let us know your Zip +4!

Non-Profit
Organization
U.S. Postage
PAID
Permit No. 361
Alexandria, VA