

Evaluating Some Popular Introductory Applied Statistics Textbooks: What We Like About Them and What We Would Like Them to Do Differently

Katherine Halvorsen¹, André Michelle Lubecke², Dexter Whittinghill³
Philip Yates⁴, John McKenzie⁴

¹Smith College, Northampton, MA

²Lander University, Greenwood, SC

³Rowan University, Glassboro, NJ 08028

⁴Saint Michael's College, Colchester, VT 05439 and Babson College, Babson Park 02457

Abstract

In this session, each panelist will respond to the following two questions for two popular textbooks and their supplements: What do you like about this textbook? What would you like this textbook to do differently? To answer these questions, each panelist will present ordered lists of 10 items, two for each textbook. By these critiques of eight different textbooks, the panelists will give their opinions on all aspects of the introductory applied statistics course. Hence, they will discuss topics (content), techniques (delivery), and technology (statistical software, datasets, course management software, applets, etc.). There will be ample time for audience participation at the end of the session. Based on a past extremely popular JSM session for introductory business and economics textbooks, this session should appeal to a large number of attendees, including instructors, authors, and publishers. It should also lead to improvements in these and other introductory applied statistics textbooks, which have an extraordinary influence on how statistics is initially presented to an increasingly large number of students.

Key Words: book reviews, exercises, statistical software, techniques, technology, topics

1. Katherine Halvorsen on Introductory Statistics with Randomization and Simulation (1st Ed) Diez, Barr, and Mine Çetinkaya-Rundel (2014)

Likes:

1. Well written, easy to read
2. Textbook FREE online in PDF for laptop or tablet. Paperback available for \$8.49 on Amazon
3. Examples and exercises use real data, in context
4. Examples and exercises emphasize checking assumptions/conditions
5. Data used in examples and exercises provided in .csv format
6. Extensive online supplements:
 - a) Labs for teaching R
 - b) Lecture slides
 - c) Supplementary text on material not in book
 - d) Interactive apps
 - e) Videos on statistical studies
 - f) Teacher forums

- g) Sample student projects
- h) Course management software

Likes to Do Differently:

1. Separate univariate from bivariate EDA
2. Add more discussion and exercises on types of data
3. Add more discussion and exercises on cases, variables, and values
4. Introduce correlation and regression as descriptive techniques early in text
5. Add more material on study design
6. Discuss study design in examples of data analyses in later chapters
7. Complete HW Answer Key (only Ch3-6 provided)
8. Number exercises in HW Answer Key with the same numbers used in text
9. Add HW exercises for probability material in appendix
10. Omit logistic regression; add more multiple regression

**2. André Michelle Lubecke on
Introductory STATISTICS exploring the world through data (2nd Ed)
Gould and Ryan (2016)**

Likes:

1. Excellent Writing: Tells the story of statistics; thoughtful ordering of topics
2. Excellent Writing: Accessible and interesting to all levels of readers
3. Excellent Writing: Patient 'mathy' reminders; clear examples of incorrect statements as well as correct statements of statistical ideas
4. Attractive: Great first impression; attractive pages
5. Interesting examples/exercises
6. Nice visuals and analogies for statistical concepts
7. Extensive set of tools for Students: Applets, Study Cards, Video solutions to 'complicated' problems, Guided Exercises, Chapter reviews, PowerPoint slides, E-text and online homework, One-semester text
8. Meaningful features throughout text: Snapshot, Details, Looking Back, Key Point, Caution
9. Reading Research Papers section
10. Excellent Instructor Resources: File of 'clicker questions', Notes for instructors, Index of applications by context

Likes to Do Differently:

1. Discussion of selecting a significance level for hypothesis tests: No reasons given for why one might use value other than $\alpha = .05$
2. Discussion of making a decision concerning hypothesis tests: Emphasis on the rule 'p-value less than .05, reject'
3. Prominence of the section on "Reading Research Papers"
4. Histogram intervals for Example 15 (Reading Electronics), Section on Comparing Means: Not the same, difficult to compare the distributions
5. Summary statistics for Example 17 (Rising College costs), Section on Comparing Means: Dependent samples BUT summary statistics are not for the differences; cannot verify given test statistic or p-value
6. Histogram scales in Figure 2.9 (comparing distributions): not the same
7. Grammatical error in Chapter 2: "...we would have got..."
8. Discussion of Comparing Means: Feels rushed compared to the rest of the text

9. Techniques for comparing means with independent or dependent samples: Separate into two sections
10. One-semester paperback textbook: Heavy and awkward to hold

**3. Dexter Whittinghill on
Introduction to the Practice of Statistics (8th Ed)
Moore, McCabe, and Craig (2014)**

Likes:

1. Modern Assumptions addressed on the spot
2. Lots of real, interesting data sets
3. Early descriptive regression
4. Early normal distributions
5. Models for regression and ANOVA
6. Objectives up front in section
7. Applets
8. Early two-way tables
9. Arrow diagram to illustrate design
10. “Look Backs”, “Cautions” & “Beyond the Basics”

Likes to Do Differently:

1. Nonparametrics in the print text
2. More examples in two-way ANOVA
3. RCBD in the printed text
4. Sub-section titles lost (size/color)
5. Bring significance/P-value guidelines back

**4. Philip Yates on
Statistics: Unlocking the Power of Data (1st Ed)
Lock, Lock, Lock Morgan, Lock, and Lock (2012)**

Likes:

1. StatKey
2. Bootstrapping and Randomization Tests
3. Essential Synthesis Chapters: Review Exercises & Projects
4. Instructor Resources: Videos Solutions & Video Tutorials
5. Datasets
6. Collection of Class Activities
7. No standard normal distribution table and no t-table!
8. Section Learning Goals at end of every section
9. P-values: informal strengths of evidence against H_0
10. Just the right amount of equations

Likes to Do Differently:

1. Include probability earlier in the book than the very last chapter
2. Essential Synthesis sections: I would assign these as readings for outside of class
3. CI's for proportions – where's the “plus 4 (AKA Agresti-Coull)” method?
4. Most of the exercises in the textbook are of the same degree of difficulty – mix it up a bit?
5. WileyPLUS – adjuncts have trouble merging their courses across campuses

6. “Multiple comparisons” is an optional section – I’d change that
7. Mention multicollinearity when talking about multiple regression
8. Not enough variation, in terms of difficulty, for book exercises
9. A few of the data sets above the heads of students: NFL Malevolent Uniforms & Penalty Yards
10. TI-83: Omit the few mentions of them

**5. Katherine Halvorsen on
Stats: Data and Models (4th Ed)
DeVeaux, Velleman, and Bock (2016)**

Likes:

1. Well written, easy to read
2. Humor adds interest
3. Initial definition: stats is a way of reasoning ... not a course about getting the right answer
4. Examples and exercises use real data, in context. Sidebars provide engaging pictures and real-data examples
5. Examples follow: “Think, Show, Tell” pattern throughout text
6. Instructor’s Guide: especially the section, “Importance of What You Don’t Say”
7. Text has instructions for eight different software packages
8. The “Ws” ... and one H (context)
9. Emphasizes distinction between an empirical distribution and a probability model
10. Extensive and excellent online materials available

Likes to Do Differently:

1. Cost of text is prohibitive \$181 (but e-book is \$99)
2. Discuss study design earlier in the book than Ch10
3. Discuss study design in examples of data analyses in later chapters
4. Introduce multiple regression as a descriptive technique earlier in the text (maybe as part of Ch 7???)
5. Add more terms to the index (e.g., “the W’s”)
6. Histogram often used to check for normality of data. Use normal scores plots as a standard part of residual analysis and for checking assumptions.
7. Size and weight of book make it unlikely students bring it to class
8. Success of book depends on students doing the extensive reading
9. Include bootstrap
10. Include randomization tests

**6. André Michelle Lubecke on
Understandable Statistics: Concepts and Methods (11th Ed)
Brase and Brase (2015)**

Likes:

1. “What does _____ tell us?” boxes
2. LOTS of good exercises and Guided Exercises
3. Exercises labels: Statistical Literacy, Interpretation, Basic Computation, Critical Thinking, Data Highlights, Expand Your Knowledge, or subject matter
4. Quotes from historical figures: Beginning of each chapter includes Kaplan, Nightingale, Einstein, Conan Doyle, ...

5. Two Gems: 'Old Faithful' data (Bimodal after earthquake) and 'Robots and Welding' example (Geometric distribution)
6. Gives a good description of confounding and lurking variables
7. Repeats definitions in later chapters when topics reoccur
8. Sufficient material for 'extra' topics: coefficient of variation, test of significance for ρ , Pearson's index for skewness
9. Formula Reference Guide labeled by Chapter
10. Technical manuals for JMP, Excel, SPSS on Student site

Likes to Do Differently:

1. First impression: Tables and formulas 'guide' inside front cover
2. Writing style: Feels like "Learn this word; now learn this word," especially in beginning chapters
3. Use of the interesting contexts in examples, exercises, Viewpoints for motivation
4. Attention to careful writing: A number of misleading/inaccurate statements throughout text
5. PowerPoint Slides: Entire paragraphs lifted from text and separated to form slides
6. Mismatched images: Content of photos does not always match contexts presented in text
7. Naming of the Critical Thinking Boxes: Contain definitions and key ideas; do not require critical thinking
8. Looking Forward feature: Intended motivation may not be realized
9. Deviations from the norm: H_1 , the Alternate hypothesis; $n \geq 5$ for normality of sample proportions; "highest/lowest" values
10. Title of Chapter 10: Contains names of distributions, not topics covered

**7. Dexter Whittinghill on
Statistics: The Art and Science of Learning from Data (3rd Ed)
Agresti and Franklin (2013)**

Likes:

1. Index of data describes them
2. Technology bits at the end of sections
3. Something for everybody at chapters end
4. "Extending & Applying the Concepts"
5. Significance guidelines for P-values (0.10-0.05-0.01)
6. Rounding rules (for the students)

Likes to Do Differently:

1. Don't try to cover everything
2. Modernize the assumptions (i.e., χ^2 -indep)
3. Normal plots should be a staple
4. Less "by-hand" computation (2-way)
5. SLR: free test of slope from test of ρ .
6. More examples in two-way ANOVA
7. RCBD in the printed text
8. Give intuitive formulas (r, slope, intercept)
9. Lose "weighted calculations" & ugly hist. boundaries
10. Make combinatorics, Poisson, ... "(Optional)"

**8. Philip Yates on
Statistics: The Art and Science of Learning from Data (3rd Ed)
Agresti and Franklin (2013)**

Likes:

1. Activities embedded in the sections of the book
2. Color coded graphs
3. Logistic Regression
4. Two Way ANOVA
5. Nonparametric Statistics
6. Written in a way to teach from book “out of order”
7. Datasets: General Social Survey examples
8. The layout of examples in sections
9. Binomial Distribution
10. Discussion of Type I and Type II errors

Likes to Do Differently:

1. TI-83 calculators: I'd get rid of them
2. Tables – especially the random number table
3. Breadth of topics – yes, this was also in my “like” section!
4. Absence of bootstrapping and randomization tests
5. Applets only on CD-ROM or MyStatLab
6. Combine Chapters 1 & 4 into one chapter about “data”
7. CI's for proportions – where's the “plus 4 (AKA Agresti-Coull)” method?
8. Relative risk: I would omit this
9. McNemar Test: I would omit this
10. Exponential Regression: I would omit this