Ten simple rules for writing statistical book reviews.

Christopher J. Lortie

Department of Biology, York University 4700 Keele St. Toronto, ON Canada. M3J IP3 PH: 416.736.2100 x 20588

and

The National Center for Ecological Analysis and Synthesis University of California Santa BarbaraSuite 300, 725 State St., Santa Barbara, CA, USA.

EMAIL: lortie@yorku.ca

URL: http://www.christopherlortie.info

Ten simple rules for writing statistical book reviews.

Christopher J. Lortie

York University and NCEAS

Abstract

Statistical books are an opportunity for accessing relatively deeper insights into statistics and software even outside the introductory classroom setting. There are however many resources available to the practitioner in addition to the traditional text model. Book reviews can thus provide a critical mechanism for the learner to assess whether the commitment to a specific book warrants the allocated time and effort. The ten simple rules format, pioneered in computational biology, was applied here to writing effective book reviews for statistics because of the breadth offerings in this domain including topical introductions, computational solutions, and theory. Learning by doing is a popular paradigm in statistics and computation, but there is still a niche for books in the pedagogy of self-taught and instruction-based learning. Primarily, these rules ensure that book reviews function as a form of short syntheses to inform and guide readers in deciding to use a specific book relative to other options for statistical challenges.

Keywords: statistics books, ten simple rules, writing reviews.

1. Introduction

Extensive resources now support the statistical programmer and analyst. The learner, reader, and general problem solver is thus faced with a choice of how to learn what is needed. This brief synthesis is not intended to be a comment or criticism on the pedagogy associated with successfully acquiring statistical and coding expertise, but there is evidence suggesting that up to 80% of coders do not read books to learn how to code (DeMarco and Lister 1999). This seems like an unfortunate statistic, but the philosophy of 'learning statistics by doing statistics' is not without merit and can be a viable approach to both introductory and expert learners alike (Smith 1998). Nonetheless, R and some other languages are both environments and quite literally languages that need to be mastered. Fluency in a written or spoken language conveys reason and semantics (Mol, Bus, and de Jong 2009), and statistical reasoning

(Smith 1998) with a corresponding representation of the associated mathematics (Widakdo 2017) can likely be secured by both doing and by reading (Utts, Sommer, Acredolo, Maher, and Matthews 2003). Different problems and topics can also require the statistical programmer to embrace a diversity of resources to illuminate a solution, and the depth required must be defined by the prior knowledge of an individual and nature of the challenge.

Many statistical texts can be a significant time commitment, and open electronic resources are abundant including vignettes and manuals in R. The decision to read a statistical programming book is thus not facile, and short syntheses, i.e. a review, of the relative merits of a specific offering can provide a critical decision tool to the potential reader. The 'ten simple simple rules' paper format was pioneered by Philip Bourne in PLOS Computational Biology (Bourne 2005), and it has proliferated to nearly 100 papers all functioning as a succinct, unique form of synthesis in itself. Sometimes extensive resources are summarized that support how to describe a focussed process or get a task done in many domains of the scientific endeavor. Of these rules papers, there have been three that address the review process including how to be an effective referee (Bourne and Korngreen 2006), how to write a literature review (Pautasso 2013), and how to write a reply paper (Simmons 2015). Many of these rules certainly support improvements in how to write a review of statistical books and should be consulted. Yet, book reviews in the Journal of Statistical Software for instance strongly suggest that the importance of this topic warrants specific treatment because these reviews can serve many functions from descriptive summary to critical analysis to launchpad for the importance of a statistical test/approach/program/language/package. All are important functions that advance statistics, but at least some of the rules here can enhance their capacity to assess merit and need for the end practitioner. (Appropriately) defend books. Write reviews. Use reviews. Book reviews that effectively support the decision process for better statistical reasoning are needed. These rules promote this paradigm shift for writing reviews.

2. Rules

Rule 1: Introduce the topic
The book title is an excellent starting point for the reader to assess whether this is the resource for her but not the only mechanism. The book cover or sleeve synopsis and publisher description can also fail to capture the whole story, and some statistical treatises, both introductory and advanced, necessarily invoke related principles and topics. As the objective expert of that specific text, an introduction to the necessity, scope, depth, and breadth of the topic in general can inform the reader on the challenges and solutions including types of data or domains of inquiry that this field examines. The goal of the first rule is to thus ensure that the reader is in right place - conceptually at least.

Rule 2: State assumed audience (i.e. expertise-level for reader)
Most technical book reviews state the level of expertise required by the reader. This is a critical form of synthesis that should be mentioned, even in brief, in a book review for statistics. The most typical categories range from introductory to advanced with relatively higher-level offering described by 'graduate student' and beyond as the reader. If the text is a blend of theory and practice with significant programming, the review should further explain the

relative expertise needed for each and if both dimensions are aligned in the assumed relative audience. Book reviews can also take the opportunity here to frame this assessment by the expertise of the referee (i.e. it is sometimes useful to know if the book reviewer is a statistician, a programmer, or a domain-specific end-user) or by the intended use of the text such as primer, guide, in-depth treatise, or textbook appropriate for instruction at a given level.

Rule 3: Explain different editions

If more than one edition, it is useful to describe the revisions to the more recent version of a book. Professional and teaching textbooks can be relatively expensive, and a critical assessment of value can provide instructors with the merits associated with potentially higher costs to students of purchasing a newer text. At minimum, a list of additions will facilitate a more informed choice for the reader and instructor and mention of case studies, updates to code and datasets, and addition of supplements are all important criteria for the choice to learn or seek solutions from a specific edition.

Rule 4: Describe the structure of chapters and associated pedagogy

Organization of the content matters to all learning (Blaich, Wise, Pascarella, and Roksa 2016), and content provides context (Grossman and Stodolsky 1995). The structure of statistical and programming texts can vary significantly. The length and complexity of chapters, use of headings and subsections within chapters, and end-of-chapter summaries are not always needed but often do no harm. Case studies, appendices, datasets, and location of supplements and supporting materials should be described. Contemporary texts in statistics are often a hybrid of print and e lectronic resource materials, and description of the extent that a text functions in this capacity can influence the decision by the reader based on her preferred modality of learning and the rationale for exploring this topic. This is also a good place to mention the different formats of the book if available in print and online.

Rule 5: Summarize content

This is the results so to speak similar to a conventional scientific publication or study report. The description should be brief, topological, and highlight the most substantive elements of the book. This component of the book review need not be unduly critical but should provide an overview of the what the text offers. Some reviews take this description of what is done to also highlight what is done best and list the most valuable chapters to the reader. This can be a useful guide to the potential reader and a means to assess expectations from the book as a whole. If there are datasets or case studies that are revisited throughout the book or across multiple chapters, the extent that the chapters connect to one another can also be summarized. Mention whether the content of the book is serialized or if chapters can be read piecemeal.

Rule 6: Critique readability

Readability is an intuitive concept. It is the ease that one can comprehend writing (Rudolph 1948; Dale and Chall 1949). Complexity in syntax, vocabulary, and sentence structure should be described in a review of a statistical book. A technical book need not be a technical challenge to read. More broadly, appeal, style, and interest are important to all but the most committed readers, and it is reasonable to assume that a book on statistics provide some sense of enthusiasm for the topic, compel the reader to think deeply, and engage one with the

challenges explored.

Rule 7: List packages and linkages to statistical concepts
Within the R statistics community, there are now over 10,000 packages that extend the base language archived on CRAN. Some statistical texts are associated with not only a single statistical program or language but with a single R package. A review of a statistical book should describe the specificity of the book, explain the extent that the book relies on single solution sets or conversely contrasts alternatives in different languages/applications/packages/libaries, and frame the programming (if provided) to general statistical theory and reasoning. At times this can be self-evident if the title of the book includes mention of the programming language or software, but the breadth of the statistics and case studies illustrated is typically not evident without review of the book.

Rule 8: Compare book to other resources
Compare and contrast. As described above, there is a wealth of both short and long-form documentation available for many open coding languages used in statistics and data wrangling. There is also an extensive opportunity to seek specific solutions through numerous forums such as Stack Overflow, Cross Validated, and Stack Exchange Mathematics. Online tutorials, blogs, carpentries, MOOCs, and webinars often provide useful, and at times, deep learning opportunities. A book review need not comprehensively list all these options and compare and contrast to the principal subject text discussed, but if there is a significant alternative to consider, it should be mentioned. Finally, there are also other books. Due diligence by the reviewer suggests that a cursory skim of related titles will facilitate a better grasp of the novelty and niche of the text in question.

Rule 9: Comment on commitment
Reading a book is a relationship. The content, style, and perspective of the author(s) becomes a shared, internalized form of knowledge in a good book. As the reviewer, it is legitimate and useful to others to mention the extent that one enjoyed the text, connected with the writing and concepts, or struggled with certain elements (i.e. comment on the quality of the relationship with the book). A review should also mention the time that the reader should set aside to read and/or fully digest the content. If the 'summarize content' rule proposed above did not mention the standout, best chapters, this is an excellent spot to describe the chapters that provided the most for your buck and should not be skipped.

Rule 10: Evaluate benefits critically
In general, it is best to be decisive in writing reviews (Bourne and Korngreen 2006). Evaluate the capacity that the book delivers on its stated goal. Accept that you are part of the review process and likely have your own, specific purpose in commiting to this text. Admit this in the review by articulating the need, success of text, and decision (or not) to use the described tools, framework, or theory. Being specific and listing criteria point-by-point is useful to editors, authors, and readers (Bourne and Korngreen 2006). Similar to the peer review process for papers, be balanced, fair, and professionally critical by mentioning both strengths and weaknesses from your perspective.

3. Summary

Reading, writing and statistics. By putting oneself on the hook for a book to take notes and annotate or further synthesize these efforts and provide a review profoundly changes how one approaches a statistical and programming text (Balajthy 1984; Tashman and Edwards 2011). Higher education in the sciences and statistics has largely done away with book reviews/reports, but application and dissemination of critical thinking in statistics in the form of reviews is a learning opportunity. Capitalize on this process particularly when using a text to solve a problem and write a review. Reviewing is a both a collegial and educational service that includes oneself as the beneficiary. The rules proposed herein for writing a book review for statistics and increasingly for the associated coding or implementation of statistics and data do not mean to imply that reading texts in this domain is a burden. On the contrary, the gratification of immersion in the structured reasoning inherent in these fields is a powerful form of literacy that merits discussion by people, for people. Recommendation algorithms certainly influence many aspects of human behavior, and a book review is a reminder to take a moment and savor the story.

Acknowledgments

References

Balajthy E (1984). "Using Student-Constructed Questions to Encourage Active Reading." *Journal of Reading*, **27**(5), 408–411. ISSN 00224103. URL http://www.jstor.org/stable/40032565.

Blaich C, Wise K, Pascarella ET, Roksa J (2016). "Instructional Clarity and Organization: It's Not New or Fancy, But It Matters." *Change: The Magazine of Higher Learning*, **48**(4), 6–13. ISSN 0009-1383. doi:10.1080/00091383.2016.1198142. URL https://doi.org/10.1080/00091383.2016.1198142.

Bourne PE (2005). "Ten simple rules for getting published." *Plos Computational Biology*, **1**(5), 341–342. ISSN 1553-734X. doi:10.1371/journal.pcbi.0010057. URL <GotoISI>://WOS:000234713100001.

Bourne PE, Korngreen A (2006). "Ten simple rules for reviewers." *Plos Computational Biology*, **2**(9), 973–974. ISSN 1553-734X. doi:10.1371/journal.pcbi.0020110. URL <GotoISI>://WOS:000240867500001.

Dale E, Chall JS (1949). "The Concept of Readability." *Elementary English*, **26**(1), 19–26. ISSN 00135968. URL http://www.jstor.org/stable/41383594.

DeMarco T, Lister T (1999). *Peopleware: productive projects and teams.* Second edition. Dorset House Publishing Company, Incorporated, UK. ISBN 0932633439.

Grossman PL, Stodolsky SS (1995). "Content as Context: The Role of School Subjects in Secondary School Teaching." *Educational Researcher*, **24**(8), 5–23. ISSN 0013-189X. doi:10.3102/0013189X024008005. URL https://doi.org/10.3102/0013189X024008005.

Mol SE, Bus AG, de Jong MT (2009). "Interactive Book Reading in Early Education: A Tool to Stimulate Print Knowledge as Well as Oral Language." *Review of Educational Research*, **79**(2), 979–1007. ISSN 0034-6543. doi:10.3102/0034654309332561. URL https://doi.org/10.3102/0034654309332561.

Pautasso M (2013). "Ten Simple Rules for Writing a Literature Review." *Plos Computational Biology*, **9**(7). ISSN 1553-7358. doi:10.1371/journal.pcbi.1003149. URL <GotoISI>://WOS:000322320200037.

Rudolph F (1948). "A new readability yardstick." *Journal of Applied Psychology*, **32**, 221–223.

Simmons MP (2015). "Ten Simple Rules for Writing a Reply Paper." *PLoS Computational Biology*, **11**(10). ISSN 1553-734X. doi:10.1371/journal.pcbi.1004536. URL <GotoISI>://WOS:000364399700064.

Smith G (1998). "Learning Statistics by Doing Statistics." *Journal of Statistics Education*, **6**(3), null–null. ISSN null. doi:10.1080/10691898.1998.11910623. URL https://doi.org/10.1080/10691898.1998.11910623.

Tashman C, Edwards K (2011). "Active reading and its discontents: the situations, problems and ideas of readers." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2927–2936. URL https://dl.acm.org/citation.cfm?id=1979376.

Utts J, Sommer B, Acredolo C, Maher MW, Matthews HR (2003). "A Study Comparing Traditional and Hybrid Internet-Based Instruction in Introductory Statistics Classes." *Journal of Statistics Education*, **11**(3), null–null. ISSN null. `doi:10.1080/10691898.2003.11910722`. URL `https://doi.org/10.1080/10691898.2003.11910722`.

Widakdo WA (2017). "Mathematical Representation Ability by Using Project Based Learning on the Topic of Statistics." *Journal of Physics: Conference Series*, **895**(1), 012055. ISSN 1742-6596. URL `http://stacks.iop.org/1742-6596/895/i=1/a=012055`.

Affiliation:

Christopher J. Lortie
Department of Biology, York University
and
National Center for Ecological Analysis and Synthesis
University of California Santa Barbara
Suite 300, 725 State St., Santa Barbara, CA, USA
E-mail: `chris@christopherlortie.info`
URL: `http://www.christopherlortie.info`