Statistica Neerlandica (1999) Vol. 53, nr. 1, pp. 68–75

Effective writing in mathematical statistics

J. S. Marron*

Department of Statistics, University of North Carolina, Chapel Hill, NC 27599-3260, USA

Care should be taken in the writing of papers in mathematical statistics for two reasons. First this enhances a paper's chances to be accepted for publication in a top journal. Second the contributions of a paper will reach a wider audience if the main ideas are easily accessible. This paper gives suggestions for improvement in two directions: presentation of mathematics and organization of papers.

Key Words and Phrases: presentation of mathematics, reviewing, writing.

1 Introduction

The mathematical statistics literature has many papers that are unnecessarily difficult to read. This paper offers ideas on how to write papers effectively. Improved readability requires effort on the part of authors, but has some immediate benefits. The first is that papers that are more readable are more likely to fare well in the reviewing process. A poorly written paper usually requires an inordinate amount of effort for reviewers to comprehend, which leads to resentment that affects the final decision. In extreme cases, reviewers will devote their energies to searching for a reason to reject a paper, to avoid the need to understand the whole thing. The second benefit for authors is that more readable papers reach a wider audience, because the main ideas are more accessible. If all papers were written with the goal of efficient communication in mind, the overall effective amount of information available to both researchers and consumers from the large current literature would be greatly increased.

This paper makes specific suggestions on two aspects of effective writing.

The first is mathematical presentation. It is argued that presentation makes a very large difference in the understanding of mathematics. The main point is:

- It is much easier to digest mathematics, especially new or unfamiliar notation, if the reader *first* understands the main idea at an intuitive level.

Section 3 shows, through a simple example, how reading can be very difficult when this principle is ignored. Then it is seen how much easier the reading becomes when this principle is adhered to. Readers who want to investigate such ideas more deeply

* marron@stat.unc.edu

are referred to STEENROD, et al. (1981), AMS (1984) and GILLMAN (1987). The second aspect of effective writing that is treated here is the organization of papers. EHRENBERG (1982) has some good ideas on this. In Section 2, some suggestions are made that are based on the premise:

- Reading will be done at a variety of different levels, so the paper should be organized in a way which facilitates the process for as many readers as possible.

In both directions, too many authors of papers in statistics write with only themselves in mind. Some papers give an impression that the author feels that only those readers willing to slog through the mathematics, i.e. work hard to figure out the main ideas, deserve to share in the main ideas of the paper. While this is probably not a conscious thought in most author's minds, many papers indicate little or no apparent regard for the reader. In particular there is a tendency to write things in the way that first comes to mind, instead of taking the trouble to make carefully considered choices from the several options usually available.

Much of what is contained in this paper is common sense when considered from the reader's viewpoint. Some of the suggestions made here are likely to be controversial. There are situations where there will be good reasons for going against some of the specifics given here. However, there should be a consensus on principles and on the need for improvement.

A part of good writing not discussed here is "style" (which concerns issues such as choice of phrasing) as it is very difficult to treat this effectively in a short paper. Also, there are many good monographs available on this, such as BARRAS (1978), the *Chicago Manual of Style* (1993), FOWLER (1996), VAN LEUNEN (1985) and HIGHAM (1993). Some other issues are briefly discussed in Section 4.

Improving readability of papers is an important issue for editors as well as authors, because there is substantial dissatisfaction with many journals that publish papers on mathematical statistics. Many journals are criticized as being "unreadable". This is a matter of serious concern, both for journals run by professional societies, which are well advised to be accessible to a wide range of members, and for commercial journals, which are competing for their share of limited library budgets.

A short summary of the suggestions made is given in Section 5.

2 Effective organization

The organizational suggestions made here are motivated by the needs of readers of papers in mathematical statistics. The wide variety of these needs is discussed in Section 2.1. Remaining subsections give specific recommendations.

2.1 Needs of the Reader

All modern researchers need to keep abreast of a very large and rapidly growing literature. This is becoming increasingly difficult. Effective writing by all could do much to ease the burden. But, there is an unfortunate tendency to curse the difficulty

of keeping up with the literature, and then turn around and contribute to the problem with one's next paper.

To see how organization of papers can help, first consider how most researchers approach the literature. A common strategy is:

1. Scan a given issue of a journal for titles (and perhaps authors), picking only a few of those for more careful study.
2. Of those chosen in 1, read the abstracts, and again select only a few of these for deeper study.
3. Of those selected in 2, read the introductions, and perhaps selected other sections, to get as much of the main ideas as possible in a short time, for example 10–20 minutes.
4. Only in cases where the reader is *directly* working in the area, and can thus justify spending a substantial amount of time, read a paper from 3 in some detail.

A well written paper takes the needs of readers at each of these levels into account simultaneously. Specific suggestions for this are given in the following sections.

2.2 Title

A good title maximizes usefulness to the reader at point 1 in Section 2.1. An effective trade off is needed between the conflicting goals of:

- Brevity
- Maximal information.

A short title means that the reader can obtain the needed information more rapidly. The title should not attempt to contain all the information that belongs in the abstract. It also need not completely classify the contents of the paper with respect to all other papers.

Of course, maximal information is important for efficient scanning of titles. Another reason that longer titles are better, is that more key words result in richer cross-listings in the *Current Index to Statistics* (1996). I suggest about one full line of type as being typically long enough to interest the desired audience, and yet not too burdensome on those who are only interested in other papers in the journal. Key words should be included which make the topic generally clear.

One way of pursuing these goals is to first start with a title, and then ask: ''would I look more deeply if I saw this title?''. If so, then ask, ''which words could be cut out without diminishing my desire to investigate more deeply?'' Even authors who are consistently good at the first step, can often improve their titles by doing this trimming at the second step.

Properly addressing these goals often means authors will have to forsake ''clever titles''. For example, plays on words can be fun for experts to read. However these often make it hard for those same experts to find the paper, so the joke can easily be wasted.

2.3 Abstract

Abstract material needs to be carefully chosen. A balance between the twin goals of brevity and maximal information content should again be carefully sought. There is room for more detail than in the title, but not enough room for all ideas covered in the paper. Make sure each "high point" is included. The paper will have a better chance in the review process if it is made clear what is done, and why it is important, since this will immediately capture the interest of the reviewer.

Any recommendations for length here must be more case dependent. Longer papers will usually need longer abstracts. However, something between 4 and 10 sentences is reasonable for most situations.

Mathematical notation is rarely useful in the abstract. Sometimes notation is introduced in an abstract, and then not used at all! Even when notation is used in the abstract, the point can usually be conveyed more efficiently in words alone.

2.4 Introduction

With the goal number 3 from Section 2.1 firmly in mind, an introduction should summarize all the main ideas. A major goal should be to indicate the main ideas of the paper as early as possible. This will help with the review process, as reviewers are likely to be kindly disposed if they understand as early as possible *why* the paper is worth reading.

In my opinion (others will disagree) the introduction should have a minimum of mathematical notation. While there are many instances where the presentation is helped by use of notation, careful thought should be given as to how much should be used. Each additional piece creates a burden on the reader, so a conscientious trade-off should be made. When notation is introduced, the main principle in the upcoming Section 3 should be kept in mind.

2.5 Figures

Figures are very useful and important parts of some papers, but there is room for improvement in this area too.

A common problem is presentation of too many figures. Some authors seem to feel that every picture they generate in a project should be included in the paper. This wastes journal space and detracts from the main points of the paper.

After each picture is made, and the ideas behind it written up, a decision as to whether or not to include it should be based on:

1. Is the picture essential to the points being made?
2. Will a simple verbal description do the same job as the picture?
3. Is this picture similar to a previous one, so that words to this effect will suffice?

Captions should be included that are sufficiently detailed so readers who are not carefully reading that section can understand the content.

The content of figures is too large a topic to deal with here. Many excellent ideas can be found in TUFTE (1983), CLEVELAND (1985), TUKEY (1990), WAINER (1990) and CLEVELAND (1993).

2.6 Conclusion section

Here I make a controversial suggestion. Many people believe a good paper is wound up with some conclusions which highlight a few of the most important lessons of the paper. If everyone were to read every part of every paper, this would be appropriate. However, in view of the way that modern researchers approach the literature, as discussed in Section 2.1, I suggest that a summary of the main points is more effective if it is in the introduction instead. It is not so elegant, since the conclusions are not properly backed up at that point. But this does have the effect of leading those who have doubts to read further and more carefully.

3 Presentation of mathematics

The main point of this section is repeated from the introduction.

- It is much easier to digest mathematics, especially new or unfamiliar notation, if the reader *first* understands the main idea at an intuitive level.

Here is an example, first stating things poorly, using only mathematics, with no explanation:

> Given $x_1, \ldots, x_n$ with n odd, let $w_1 = \min\{x_i : i = 1, \ldots, n\}$, and then recursively define for $j = 1, \ldots, n$,
>
> $$w_j = \min(\{x_i : i = 1, \ldots, n\} \setminus \{w_1, \ldots, w_{j-1}\})$$
>
> where $\setminus$ denotes deletion of the values on the right from the data set on the left, and use these to define for $k = 1, \ldots, n$, $u_k = |w_k - w_{(n+1)/2}|$, and then let $z_1 = \min\{u_k : k = 1, \ldots, n\}$, and then recursively define for $l = 1, \ldots, n$,
>
> $$z_l = \min(\{u_k : k = 1, \ldots, n\} \setminus \{z_1, \ldots, z_{l-1}\})$$
>
> and then let $a = z_{(n+1)/2}$.

That was an intentionally clumsy introduction (featuring intentionally unclear and non-mnemonic notation) to the concept of the median absolute deviation from the median. However, it illustrates something which happens frequently: an idea is presented in the first way that comes to mind, with no choice made among the many ways that the concept could be presented.

Next the same idea is presented, using the same clumsy set of ideas, and also the same non-mnemonic notation, in a way that is much easier to digest simply because each bit of mathematics is preceded by a short indication *in words*, of what the author (and thus the reader) has in mind:

> The concept of *median absolute deviation* of a set of numbers $x_1, \ldots, x_n$, of odd cardinality n is based on taking the median of the *deviations*, which are the set of distances from each point to the median. The median is the central *order statistic*.

Order statistics are a relabeling of $x_1, \ldots, x_n$, in increasing order, which can be defined as: $w_1 = \min\{x_i : i = 1, \ldots, n\}$, and then recursively for $j = 1, \ldots, n$,

$$w_j = \min(\{x_i : i = 1, \ldots, n\} \setminus \{w_1, \ldots, w_{j-1}\})$$

where $\setminus$ denotes deletion of the values on the right from the data set on the left. Since n is odd, the index of the central order statistic is $(n + 1)/2$, so the median is given by $w_{(n+1)/2}$. Next define the deviations, for $k = 1, \ldots, n$, $u_k = |w_k - w_{(n+1)/2}|$. Now again take the median, but this time of the deviations, $u_1, \ldots, u_n$. The corresponding order statistics, defined in the same recursive fashion as above, are $z_1 = \min\{u_k : k = 1, \ldots, n\}$ and for $l = 1, \ldots, n$,

$$z_l = \min(\{u_k : k = 1, \ldots, n\} \setminus \{z_1, \ldots, z_{l+1}\})$$

The median absolute deviation is then the median of these, given by $a = z_{(n+1)/2}$.

Note that the second version is much easier to digest, because there is no need to individually figure out the idea behind each piece of notation, since it is explained first in each case. Now the same concept is presented, using a better overall approach, and more mnemonic notation. It may be possible to improve this further, but the point here is how much easier it is to digest this than the first attempt.

The *median* is a key concept underlying the *median absolute deviation* of a set of numbers. The median of $x_1, \ldots, x_n$, when the cardinality n is odd, is the central *order statistic*. Order statistics are a relabeling of $x_1, \ldots, x_n$, in increasing order, say

$$x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$$

so that $x_{(i)}$ is the ith smallest of the x_i's. Since n is odd, the subscript of the central order statistic is $((n + 1)/2)$, so the median is given by $m = x_{((n+1)/2)}$. The median absolute deviation of $x_1, \ldots, x_n$ involves two applications of this operation. First find the *center point* of the numbers, given by m above. Then to measure *spread*, take the median of the set of distances from each point to the center, called the *deviations*. These are given by $d_i = |x_i - m|$, $i = 1, \ldots, m$. The median absolute deviation is then the median of this set of numbers, $MAD = d_{((n+1)/2)}$.

Such differences in presentation make a huge difference in how much time is required for reading two papers which have exactly the same content, but are written differently. The most important part is *motivation*. Mathematics are easier to digest when the reader properly understands the concept, and why he should be digesting it. There is some cost in terms of added length, but time on the part of researchers should have a higher priority than journal space, especially for journals that aspire to be ''readable''.

4 Points deliberately omitted

Some points that might be considered in a discussion of "effective writing" have been deliberately omitted here. These include:

1. Length of papers. This must necessarily depend on the topic being addressed. However, it is important to realize that shorter papers have a much higher chance of acceptance, mostly because referees feel much more kindly disposed to them (think of your own feelings when asked to a review a paper of more than about 25 pages).
2. The proportion of text to mathematics. This must also depend on the context.
3. The "I" vs. "we" issue. I have heard this hotly debated, and find something to both sides. In my opinion, choice of which to use is a personal matter. Authors should choose whichever they feel most comfortable with, as this stimulates creativity and allows more thought to be devoted to more important aspects of effective writing. However personal pronouns appearing too frequently can be distracting, and should be avoided.

5 Summary of recommendations

- Title: 1 full line.
- Abstract: 4–10 sentences, no mathematical notation.
- Introduction: words only.
- Conclusions: not needed, since main ideas should be in the introduction.
- Presenting mathematics: explain ideas behind notation *first*.

6 Acknowledgement

The writing of this paper was supported by NSF Grant DMS-9203135. Many students at the University of North Carolina have provided helpful input. R. J. Carroll, P. Hall, R. Kohn, I. Olkin and M. P. Wand made useful comments. R. Boyce helped find monographs on good writing. Anonymous reviewers have made many helpful suggestions. One reviewer seemed to enjoy making it clear that my own "style" leaves much to be desired!

References

AMERICAN MATHEMATICAL SOCIETY (1984), *Manual for authors of mathematical papers*, American Mathematical Society, Providence.

BARRAS, R. (1978), *Scientists must write: a guide to better writing for scientists, engineers and students*, Chapman and Hall, London.

The Chicago manual of style (1993), University of Chicago Press, Chicago.

CLEVELAND, W. S. (1985), *The elements of graphing data*, Wadsworth, Belmont, California.

CLEVELAND, W. S. (1993), *Visualizing data*, Hobart Press, Summit, New Jersey.

Current Index to Statistics (1996), The American Statistical Association and the Institute of Mathematical statistics, Alexandria, Virgina.

EHRENBERG, A. S. C. (1982), Writing technical papers or reports, *The American Statistician*, **36**, 326–329.
FOWLER, H. W. (1996), *The new Fowlers's modern English usage*, Oxford University Press, Oxford.
GILLMAN, L. (1987), *Writing mathematics well: a manual for authors*, Mathematical Association of America, New York.
HIGHAM, N. J. (1993), *Handbook of writing for the mathematical sciences*, Society for Industrial and Applied Mathematics, Philadelphia.
STEENROD, N. E., et al. (1981), *How to write mathematics*, American Mathematical Society, Providence.
TUFTE, E. R. (1983), *The visual display of qualitative information*, Graphics Press, Cheshire, Connecticut.
TUKEY, J. W. (1990), Data-based graphics: visual display in the decades to come, *Statistical Science*, **5**, 327–339.
VAN LUENEN, M. C. (1985), *Handbook for scholars*, A. A. Knopf, New York.
WAINER, H. (1990), Graphical visions from William Playfair to John Tukey, *Statistical Science*, **5**, 340–346.

Received: February 1996. Revised: June 1997.