A Checklist for Preparing Your Manuscript

Jacob Bien

(Last modified: June 8, 2022)

When you're preparing a manuscript for submission to a journal, the journal will often have a list of stylistic guidelines that you must follow. However, much before you're ready to submit to a journal, you will be preparing a manuscript for me to read and give feedback on. Before you send me your manuscript, I would ask that you please start by making sure that your manuscript adheres to the following guidelines. Some of these items may seem obvious to you while others may seem less obvious. I assembled this list based on many years of editing student manuscripts.

- ☐ Being consistent
  - ☐ Do all titles, section titles, and subsection titles have a consistent choice of how to capitalize? The two standard choices would be "My Wonderful Title"[1] and "My wonderful title." I am ok with either of these conventions as long as you are consistent throughout the paper.
  - ☐ Is it "dataset" or "data set?" I'm fine with either but it's important to be consistent through the paper.
  - ☐ Are you using present tense or past tense in your simulation study and data analysis sections? Whichever tense you choose, use it consistently.
- ☐ Capitalization
  - ☐ Terminology is all lower case unless it is based on someone's name. So "elastic net" (not "Elastic Net") but "Wasserstein distance" and "Gaussian process."
  - ☐ Notice the difference in capitalization when you write "in the next section" versus "in Section 3." Likewise you could say, "in the appendix, we show..." whereas you would write "in Appendix A, we show...." The same goes with theorems and figures.
- ☐ Latex equations
  - ☐ Treat equations as part of the sentence. Example: "I like $A = \pi r^2$." Notice that you could replace the equation with a word like "food" and it would make sense.
  - ☐ After a display equation you should almost always have either a period or a comma. As in, "My favorite model is
$$y \sim N(X\beta, \sigma^2 I_n).$$"

[1] A nice tool for understanding the conventions for capitalization in titles is `https://titlecaseconverter.com/`. I recommend checking the "Show Explanations" option.

Right before a display equation, you would only use a colon if it is ending a complete sentence. So you wouldn't say, "We write this as:" since that's not a full sentence. But you could say, "We write this as follows:" An alternative would be to write, "We write this as" without any comma.

- ☐ Don't start sentences with mathematical notation. So instead of writing, "Σ is a Toeplitz matrix," you would find some other way to word it. For example, you could write, "The matrix Σ is Toeplitz."
- ☐ Use `$\mathrm{diag}$` rather than `$diag$` (and likewise for other words).
- ☐ Be sure to write `$p$` rather than just `p` in text (and likewise for other cases where a single-letter variable name appears within a line of text). The reason this matters is that the result will look different (p versus p).
- ☐ For inner products, use `$\langle x, y\rangle$` rather than `$<x, y>$`. This will results in $\langle x, y\rangle$ rather than $< x, y >$.
- ☐ For "much greater than," use `$p \gg n$` rather than `$p >> n$` (this will result in $p \gg n$ rather than $p >> n$). And likewise, use `$\ll$` for "much smaller than."
- ☐ For norms, use `$\|$` rather than `$||$`.
- ☐ Use `\left` and `\right` whenever using parentheses, curly braces, etc. This will resize them as appropriate for the context.
- ☐ Refer to equations either as "Equation 12" or "(12)," but not as "Equation (12)." Use `\eqref{}` for references of the form "(12)."

☐ Citations

- ☐ Use `\citep{}` for a parenthetical and `\citet{}` for a noun. Examples: (a) "The lasso `\citep{tibs96}` uses an L1 penalty." (b) "`\citet{tibs96}` shows that..." One places a citation in parentheses as in (a) when you could read the sentence without the citation and it would still make sense. This is not true in (b), where the citation is the subject of the sentence.
- ☐ Be sure that you don't have double parentheses. To do so, you can use `\citealt{}`. Example: "The lasso (see, e.g., `\citealt{ESL}`) is..."

☐ Common sense things

- ☐ Is every sentence grammatically correct?
- ☐ Are all words spelled correctly?

☐ Logical flow

- ☐ When a sentence refers to something as if it has already been introduced to the reader, has this thing actually already been introduced to the reader? [The only way I know to figure this out is to read your writing from the very start to the very end. And as you're reading, pretend you're someone who is reading this paper for the first time.]
- ☐ Whenever referring to a piece of notation or a fact that was used more than, say, two paragraphs ago, provide an equation number or section reference for where this notation or fact was presented. E.g. "since $a + b = c$ (see Theorem 2)" or "Recall the definition of W from Section 2."

- ☐ Prepare the reader for a theorem before presenting it (and likewise propositions and similar). Also, sections should rarely end with a theorem; rather, they should end with some text either explaining something about the theorem or transitioning the focus to the next section.
- ☐ Any technical claim should have an explanation. This can be a reference to a past equation, a citation to an external reference, or a reason (if the reason can't fit in a sentence or two then the technical claim should be its own result, with proof in appendix).

☐ Plagiarism: All writing must be your own. There should absolutely never be verbatim (i.e., word-for-word) phrases taken from another reference unless you put them in quotes with attribution. This is an academic integrity issue, and it is problematic even to copy phrases from your **own** writing (this is called self-plagiarism). I have seen this come up a surprising number of times in statistics papers (even published ones!), especially in the context of describing data sets and describing related work. If you need to describe a data set or someone else's method, I would suggest the following process: Read what they wrote and then wait some time (maybe a day?) and then write the description yourself *without looking at the original text*. This will make it more likely that you are describing the idea in your own words. If there are a lot of details that are hard for you to remember without looking back at the source, this may indicate that you should instead be referring the reader to read the description in that paper. Or if you really do need the details in your own paper, that might be a situation where a direct quote with attribution is appropriate.

Thanks to Daniela Witten for contributing some of these!