

Some Recommendations for Authors Using LaTeX

Dr. Mirko Janc

Technical Typesetting Inc.

1510 Caton Center Drive, Suite E, Baltimore, MD 21227

Phone: (410) 247-8717; E-mail: tti@clark.net

26 October 2000

Introduction. For submission of TeX manuscripts to be published with the American Statistical Association we strongly recommend use of LaTeX with its standard ARTICLE style. No special esoteric additional styles should be used unless absolutely necessary, e.g., to include Encapsulated PostScript graphic files, to set commutative diagrams, etc. Private macros should be used sparingly, and they are acceptable only in the form of abbreviations of longer strings. Only macros used in the manuscript should be supplied with the article (not the whole macro libraries in use at the author's site).

Some of the standard LaTeX features do not work properly from the point of view of the final typeset product, so we give in this text only a short list of most important hints. Following these hints will help the American Statistical Association to use author's key files effectively.

One of the important things to bear in mind is that the width of the typeset column is small (21 pica at 9pt point size) compared to the pages produced with standard LaTeX styles, so if options 10pt, 11pt, or 12pt are used, then the `textwidth` should be set to 23 pc, 25 pc, or 27 pc, respectively. This setting will fit the long formulas to the proper size, keep them from being rebroken by the typesetter, and reduce the possibility of introducing new errors in the production process.

References and Numbered Environments. References should be supplied in the standard LaTeX format using `\bibitem` and `\thebibliography` environment. Supplying the bibliography in a form so that at production stage BibTeX must be used to draw the necessary items from the database is not acceptable. References should conform to the basic American Statistical Association style.

No private special counters should be used to generate the style of citations and references, as well as that of automatically generated `theorem`, `lemma`, `proposition`, `definition`, `example`, etc., environments. If theorems and propositions are needed, a declaration in the preamble such as

```
\newtheorem{theorem}{Theorem}
\newtheorem{proposition}[theorem]{Proposition}
```

or

```
\newtheorem{theorem}{Theorem}
\newtheorem{proposition}{Proposition},
```

depending on the numbering style, is all that is needed to get the standard LaTeX version of the theorem environment.

Sizing of Fences. Standard LaTeX supplies several standardized sizes of fences:

$$(), \quad (), \quad (), \quad (), \quad ().$$

The standard way to key them is, in that order, `(,), \bigl(, \bigr), \Bigl(, \Bigr), \biggl(, \biggr), \Biggl(, \Biggr)`. For correct sizing in most cases the basic two are enough:

$$() \text{ for } () \quad \text{and } \biggl(, \biggr) \text{ for } \left(, \right).$$

In most one-line formulas, only the standard size of fences should be used. The next size (1.2 times larger) should be used sparingly, only in cases of heavy sub- and superscripts, or as outer fences to (technically) complex terms already in standard size fences. The use of self-sizing `\left(, \right)` should be avoided because they tend to cover the full height of the formula. Mathematically speaking, their size is very unstable (and discontinuous—being a step function) with respect to the height of the subformula, and it tends to render oversized fences that are not esthetically pleasing.

The `\bigg` size should be used to cover most of the expressions with fractions, integrals, summations, etc., in displayed formulas. The `\Big` and `\Bigg` sizes cover 1.5 and 2.5 lines, respectively, and in cases where they are really needed, their use is preferred to `\left, \right`.

The use of `\left, \right` is appropriate in cases where the fences have to be considerably bigger, such as in matrices, “cases,” etc.

Displayed Formulas. The main (technical) concern in the case of displayed equations is to break them properly to fit in the width of the journal. For multi-line formulas, the most often used LaTeX constructions are `eqnarray` and `array` within `equation`. Whereas in the `eqnarray` environment all of the alignment fields are set in display style (they automatically build up fractions, use the bigger version of operators with limits above and below, etc.), in `array` the default is text style, so to generate the built-up version, the contents of the field should be preceded by `\displaystyle` (which, of course, could be abbreviated to something like `\DS`).

Some Small Hints.

Delimiters. Parentheses, brackets, and braces in LaTeX are set to act as opening and closing delimiters, so their correct use poses no problem. If braces are unmatched—we suggest the use of `\lbrace` or `\rbrace`—the longer version of `\{` and `\}`—because these codes make locating unclosed groups easier when the file is debugged.

`\mid` and `\parallel` are often misused for absolute values and norms. The symbols look the same, but `\mid` and `\parallel` have additional space before and after (acting as binary relations). The angle brackets should not appear as `< >`, but as `\langle \rangle`. Their proper names are `\langle` and `\rangle`, and they also can be sized when prefixed by `\big` codes.

Commas and Spacing. In expressions containing differentials, a thin space `\,`, should appear before the differential as in $\int f_n(x) d\mu(x)$. Another common mistake that leads to poor spacing and line breaks is the use of a comma with conditions within the paragraph. Formulas such as $2^{2^n} + 1, n \in N$, should be keyed as `$2^{2^n}+1$, $n\in N$`, not as `$2^{2^n}+1, n\in N$`. In the latter case the space after the comma would be too small and would prohibit a line break at that place, moving the possible break to a less pleasing position.

The comma used in math mode is set up so that it automatically generates a thin space after it, which is in most cases what is needed when enumerating cases, elements of a set, etc. No `\,` should be added after the comma in these cases. `\{a,b,c\}` gives $\{a,b,c\}$; omitting the `\,` codes will not lead to $\{a,b,c\}$ as less experienced LaTeX users probably suspect when deciding to key the thin spaces. On the other hand, a number such as “\$10,500,000” would look like “\$10,500,000,” if it appears in math; these spaces can be avoided by putting the comma in a pair of braces to make it an ordinary character and not an intelligently acting math punctuation sign.

Special Characters. Openface letters (which in extended standard TeX distribution reside in `msbm*` fonts in the additional collection from the American Mathematical Society) should be keyed as `\Bbb R` and not, for the openface \mathbb{R} , as `\IR` (with or without some kerning between \mathbb{I} and \mathbb{R}).

The German “Fraktur” characters should be keyed as `\frak X`, and they must be borrowed from the Euler collection from the American Mathematical Society (fonts `euf*`).

Unless there is a strong reason to do so, the script ℓ (`\ell`) should not be used instead of italic l (`\l`). The style prefers *letters* to *elletters*, and the relation $k, l, m \in N$ looks better than $k, \ell, m \in N$.

New Mathematical Operators. For mathematical operators that are not covered in standard LaTeX, use the Plain-T_EX definitions such as

```
\def\liminf{\mathop{\rm lim}\nolimits}\limits}%
\def\Conv{\mathop{\rm Conv}\nolimits}%
\def\Var{\mathop{\rm Var}\nolimits}%
\def\Wlog{\mathop{\rm wlog}\nolimits}%
```

to get

$$\int_E \liminf_{n \rightarrow \infty} f_n d\mu \geq \liminf_{n \rightarrow \infty} \int_E f_n d\mu \quad \text{or} \quad \text{Conv } X = \{f \mid f \text{ is convex on } X\}.$$

etc. Notice that ‘wlog’ was called `\Wlog` because `\wlog` is a reserved TeX command.

Large Formulas. The introduction of intermediary denotation for long clusters of subformulas sets long expressions in a more compact form. Such an approach is nearly always a must with matrices that cannot fit into the width of the journal, but it may be useful also in “ordinary” formulas. A subformula

$$P\left\{\max_{0 \leq k < m/m_0} \max_{km_0 \leq i < (k+1)m_0, (k+1)m_0 \leq j \leq (k+1)m_0} |S_j - S_i|/(j-i)^{1/2} > b\right\}$$

of a considerably larger display could be rewritten also as

$$P\left\{\max_{0 \leq k < m/m_0} \max_{A_k(i,j)} \frac{|S_j - S_i|}{(j-i)^{1/2}} > b\right\}$$

where $A_k(i, j): km_0 \leq i < (k+1)m_0 \wedge (k+1)m_0 \leq j \leq (k+1)m_0$. Of course, this is only one of many possible ways to shorten the expression above. This is only an example of the possibility to rewrite formulas that are technically too complex.

Tables and Figures. The style of tables in all publications of the American Statistical Association is quite specific and not easy to set for “non-professional” typesetters. They use a special sans-serif font setting (including formulas), which is unavailable in Computer Modern fonts that come with standard TeX implementations, as well as conform to some other guidelines/rules that can not be controled from a LaTeX style. Because of that, the most useful way to supply tables (as a source on which to implement the style elements) is to key them in the standard LaTeX `tabular` environment without using any sophisticated formatting. Use `\caption{...}` before the table itself.

For figures, the `\caption` comes after the `\vspace*{...}` or something like `\epsffile{...}` (for `.eps`, or `.ps` figures).