

MT4613 STATISTICAL RESEARCH IN PRACTICE

GUIDANCE ON WRITING A STATISTICAL REPORT

Prepared by: Len Thomas (len@mcs.st-and.ac.uk), University of St Andrews

Last updated: 2 November 2010

There are many ways to write a good statistical report, but even more to write a bad one! The following advice will hopefully help you in preparing project reports in MT4613, and also be useful if you have to write reports or scientific manuscripts for other courses or in your future career. Feel free to use these tips selectively – you don't have to slavishly follow them all. However, if you ignore some please have a good reason for doing so!

GENERAL ADVICE

A statistical report is just a particular type of scientific report, and so almost all of the advice about how to write a good scientific report or paper is relevant. There are tons of books on this, for example some good ones are Porush (1995) and Day (1994). Of course, a scientific report is just a particular type of written document, so most of the general advice you can find about writing clearly is also relevant. Particularly good examples are Strunk (1918) and Williams (2009). There is also lots of advice available on the internet about good statistical, scientific and general writing.

- *Consider your audience.* Never write a document without first considering who it is intended to be read by. For MT4613, you should assume that your report is intended for a professional scientist who is not a statistician by profession, but who has a reasonable grasp of basic general statistics. You should assume they do not know about the particular problem you are writing about. Hence, you will need to explain the background to the problem, and give enough details about the methods used so that they can follow what you did. You don't need to give details of standard methods, such as calculation of confidence intervals using a normal approximation – but you should consider whether to refer to a standard text for such things, for example “Confidence intervals were calculated using the normal approximation to the binomial distribution (Minderbender 1908).” (OK, I made that reference up!) An excellent skill to develop is the ability to explain what you did in enough detail that a naïve reader can follow it quite easily, even if what you did was quite complex.
- *Read some papers for inspiration.* If you've never read any statistical literature, then head over to the maths library and spend some time flicking through the journals. This should give you a good feel for how academic papers in statistics are structured, and the variations in style. Don't worry if you don't understand the detail – it's the general layout and style you should examine. The first case study in MT4613 involves a critique of what is effectively a statistical report, so this should also help.
- *Acknowledge your sources.* If you use material or ideas from others, acknowledge this in the report. If you're building on previous published work, refer to that work in the text of your report (e.g., “Methods for the analysis of catch-at-age data are described by Minderbender (1908). In summary [summarize in a few sentences]. Here, we [describe how you extend them, if you do]”). If you got some help from someone in writing the report, include an Acknowledgements section, and thank them there (see Acknowledgements, below).
- *Be careful which sources you use.* It's better to refer to published work than unpublished, if possible – so try not to refer to lecture notes, for example, if the notes are actually based on something that's published. Refer to the published original material instead. Avoid referring

to web sources unless you can't find anything in a peer-reviewed journal article or text book that says the same thing. (Web resources are notoriously unreliable, and most are not considered kosher sources of reference material by scientists.)

- *Keep it short!* Longer reports are not necessarily better reports – in fact they are often worse as they contain lots of unnecessary material, or repetition. In business, time is money, and overly long writing will not be read. In academia, there are often stringent page limits on publications. In MT4613, instructors are human, and will not relish reading verbose reports! Therefore, only include material that helps the reader understand the problem you tackled, what you did, what you found and what you conclude from these findings. Be concise.
- *Practice makes perfect.* It takes quite a bit of effort to develop a good style. Use feedback from lecturers to improve.

STRUCTURE AND CONTENT

Scientific reports tend to have the following structure: abstract, introduction, methods, results, discussion, conclusion, acknowledgements and references. Each element is often a separate section, although sometimes sections are combined (e.g., discussion and conclusion), and in some journals the order is changed (e.g., for the journal *Nature*, the methods section is put in small type at the end, accurately reflecting the importance the editors of that journal ascribe to rigorous scientific method). Very technical material, extensive tables of results, and other non-essential material are often relegated to appendices – although you won't need to do that in your reports for MT4613. Often, articles do not follow this structure at all – for example review articles – so you should feel free to be flexible if you think your material warrants it.

- **Abstract.** (This section is also called “Executive summary” in business reports, and is sometimes called “Summary” in scientific publications.) It should summarize the background, methods and main findings of the paper. Someone reading this should be able to get the gist of what you did and what you found, and can use it as the basis for deciding whether to read the main report. Keep it very short (perhaps 250-350 words max; even shorter if you can manage it).
- **Introduction.** Give a concise background to the problem, referring to previous work in the field if that's appropriate. For MT4613 you need to demonstrate that you understand the problem. Give a motivation for the work that's going to be described in later sections – why did you do it (i.e., why is it scientifically worthwhile/interesting)? “Proper” scientific papers may also lay out the scientific hypotheses being tested – this might not be applicable to the reports you write for MT4613. Finish by describing in a couple of sentences how the rest of the report is structured (especially if you use a non-standard layout).
- **Methods.** Describe what you did, in enough detail that someone else could reproduce it. In academic papers this is sometimes not possible due to space constraints, so full details are relegated to an appendix; sometimes the computer code and input data are archived in an appendix (especially online appendices). For the reports in MT4613 you should be able to explain what you did concisely without the need for appendices. If the methods section is quite long, you may need subheadings to divide it into manageable parts. Don't be afraid of using formulae here to explain your methods, but make sure you define all the symbols used. It may also be helpful to use a figure, for example to show a map of the study area, or a diagram explaining how various analyses fit together. Such figures will likely not be required in MT4613.
- **Results.** Say what you found. Results should be given in the same order as the work was laid out in the methods section, with the same (or nearly the same) subheadings as the methods

(if you had any). Only report results that are relevant to the problem at hand, and were motivated by the methods section – don't just dump any old output from a computer program into the results section. Summarize results into tables and figures as much as possible, and don't repeat information given in tables in the text, just refer to it. For example "All five populations showed a statistically significant decline in numbers after the fire (Table 1)." See below for more advice on tables and figures. Consider how many significant figures/decimal places results should reasonably be reported to, and be consistent – there is often no need to report results to more than 2 or 3 decimal places.

- Discussion. This section usually starts by discussing how reliable the results are. Lay out the assumptions of the methods used, consider how valid they are, and what the consequences of likely violations might be. After this, the discussion considers the implications of the results – what do they mean for the questions being addressed by the paper. This section often finishes with a brief consideration of what further work might be useful, often with specific suggestions.
- Conclusion. Concisely lay out the main conclusions of the paper. This section is often somewhat redundant, given that these are also given in the abstract (although in less detail). You're probably better to finish the discussion with a pithy concluding paragraph.
- Acknowledgements. You may not need this section.
- References. This section is sometimes called "Literature cited". Each printed text or web site referred to in the report should be listed here, in a standard format. The format can mimic any standard statistics journal; if in doubt, use the Harvard system of referencing (search on the web for information about this), which is something of a standard in science.

STYLE

In the good old days when I were a nipper, scientific writing style was very prescribed. One used the passive voice and past tense. Now, things are less rigid. Nevertheless, you should aim for a clear, precise presentation style and stick as much as possible to whatever voice and tense you choose.

- The methods and results sections should be in the past tense, as you are describing what you've done and what you found. The introduction is often in the past tense when describing previous work, and present tense when describing the ideas addressed in the report and the report layout. Discussion and conclusions are often in the present tense. However, as I said, this is somewhat fluid. Look at some published papers for guidance.
- Most journals now prefer the first person plural active voice (e.g., "Here, we show that analysis of too much data is bad for students' brains."), even if the report has a single author. If you want to stick to the passive voice (e.g., "Analysis of too much data was shown to be bad for students' brains.") then that's fine.¹
- Keep sentences short; don't use long words unnecessarily; use correct English; avoid slang and jargon; spell out all acronyms the first time you use them. Basically, write as clearly as possible.

¹ A colleague pointed out that in his day, when he were a nipper, data had to be plural – e.g. "too many data" rather than "too much data". I think it's now deemed acceptable for data to be singular, but opinion clearly varies on this.

TABLES AND FIGURES

- Figure and table captions should, as a rule, contain enough information that a qualified reader can understand the figure or table having read only the caption. Everything in the figure or table should be self-evident, or explained in the caption.
- If you present figures or tables, they should be relevant and you should say what their relevance is. Do not just dump output from some computer package into the report.

MISCELLANEOUS ADVICE

- For MT4613 you need to demonstrate that you understand the methods you're using. Make sure you justify your choice of analysis method (e.g., in the Discussion), and perhaps, if appropriate, briefly mention alternatives and what their strengths and weaknesses might be.
- If you fit a model to some data, consider whether there is a need for goodness-of-fit tests. State the assumptions of the methods you use, and consider how likely they are to be met – see the above advice on the Discussion section for more on this.
- Leave enough time to write a decent report. Good work badly reported will gain few marks.

ACKNOWLEDGEMENTS

This document is partially based on Rootzén (2003). Thanks also to David Borchers, Steve Buckland and Janine Illian for suggestions. Any further comments/suggestions would be appreciated.

REFERENCES/USEFUL ADDITIONAL MATERIAL

Day, R.A. 1994. How to Write and Publish a Scientific Paper. Oryx Press.

Porush, D.A. 1995. A Short Guide to Writing About Science. Harper Collins.

Rootzén, H. 2003. How to write a statistical report. Unpublished document, dated 15/1/2003. Available at <http://www.math.chalmers.se/Stat/Grundutb/CTH/tms061/0809/reportwriting0315.pdf> Accessed 20/10/2010.

Strunk, W. Jr. 1918. The Elements of Style. Press of W.P. Humphrey.

[Yes, this book really was first published in 1918, and you can find the original edition online. However, good writing style has changed a little over the intervening decades, so it's probably better to look at one of the more modern editions – for example, the fourth edition, published in 1999.]

Williams, J. M. 2007. Style: Lessons in clarity and grace (9th edition). New York: Pearson Longman.