PH3010 / Advanced Skills

Introduction to Statistical Data Analysis

Week 3: Report Writing

Autumn term 2019

Glen D. Cowan
RHUL Physics

Aims of this talk

This talk is part of the Department's broader programme on writing skills. For now its focus is on PH3010, but most of what is discussed is equally relevant to any similar paper (e.g., your final-year project reports).

There are many issues to keep in mind when writing a report. For now we will look at just a few key elements:

Communicating to the right audience at the right level.

Keeping your message organised in a clear and logical way.

Good style, grammar, punctuation, etc.

The audience

When writing anything, consider carefully:

Who is going to read it?

What is their prior knowledge of the subject?

For your project reports, take your target audience to be physicists who are not necessarily specialists in its specific topic.

Try to keep the level of the report consistent. Do not start with

All matter is made up of tiny particles called atoms.

and on the next page say

At zero detuning, paths through adjacent AT split states interfere constructively leading to enhanced absorption EIA on the three photon resonance $|0\rangle \rightarrow |1\rangle \rightarrow |2\rangle \rightarrow \{|3'\rangle, |3\rangle\}$.

(Kondo et al., arxiv:1510.01729)

What your reader does and doesn't know

Consider two ways of describing part of an analysis:

(1) An estimate of θ was obtained.

(2) The analysis resulted in an estimate of the asymmetry parameter θ, which is defined through the probability density given in Eq. (7).

With (1), the reader probably thinks something has been missed and may try to skip back to find where θ was introduced.

In (2), it could be that this quantity is being mentioned for the first time and it is not expected that the reader should have prior knowledge about it.

Adjusting sentences to fit the context

1) The analysis used the profile likelihood.

2) The analysis used a function called the profile likelihood, which only depends on the parameters of interest.

3) The analysis used the profile likelihood as defined in Eq. (12).

4) The analysis was based on a function called the profile likelihood, obtained by evaluating the likelihood with constrained estimates for a subset of its parameters (see, e.g., Ref. [37]).

The choice of the most appropriate sentence will depend on the context. Ask yourself at the point where the sentence appears what the reader knows about the profile likelihood and what role this information plays in what you're trying to say.

The chapter/section outline

A possible (incomplete) outline:

1. Introduction
2. Mathematical background on statistical methods and the method of least squares (here OK to combine Sec. 1 and 2)
3. Exercise 1: Polynomial fits with the method of least squares
   - 3.1 Statement of the problem and methods used
   - 3.2 Fit results
   - 3.3 Error propagation
4. Exercise 2: Analysis of Galileo's ramp data
   - 4.1 ...
5. Exercise 3: Analysis of Ptolemy's refraction data
   - 5.1 ...
6. Summary and conclusions

The introduction

The introduction should cover two main points:

First, it should state briefly what the report is all about and mention some motivation for why the topic is being covered.

Second, it should give a sort of "roadmap" to the rest of the report. This can be of the form:

"Section 2 provides background on the statistical methods used, with focus on the method of least squares. In Sections 3, 4 and 5, these methods are applied to problems of..."

The roadmap is not always included in published books and papers but we suggest doing this in your reports.

Section on background information

Most longer reports that you will write should have a chapter or section that reviews some mathematical or theoretical background that is necessary to understand the subsequent parts of the project.

For, e.g., a final-year project this could include an overview of the theory needed to understand the result of an experiment you have carried out, and if relevant something about its context, e.g., with respect to previous results.

For purposes of the module on Statistical Methods, this section can simply be a brief summary of the key elements on statistics and curve fitting covered in the script. Its purpose is to make the report more self-contained

Alternatively, given the nature of this mini-project, some brief background can be given in the Introduction and further info could be placed into the sections corresponding to the different exercises.

Structure of chapters and sections

More relevant for longer reports with a full chapter structure; here optional but good if you have many subsections.

In analogy with the introduction, each chapter should start by explaining its place in the larger narrative and providing a roadmap to its sections, e.g.,

Chapter 5

Theoretical interpretation of recent measurements

The experimental results reviewed in Ch. 4 have provided the basis for a number of tests of theoretical models. Section 5.1 summarises tests of supersymmetric models based on the measured Higgs boson production rates. Higgs boson mass measurements have been used to test a variety of models such as compositeness, as described in Sec. 5.2.

5.1 Tests of supersymmetric models

...

Sections and paragraphs

If you find yourself using a concept in a section that is only defined later on, consider reordering.

If you find yourself describing something in a section that is unrelated to the section title, consider whether it could be discussed better elsewhere, or whether the section could be broken into subsections.

Each paragraph should address a single well-defined topic. If you find yourself talking about different issues in different parts of a paragraph, it's too long.

Before you dive into a detailed technical description, make sure your reader understands why this detail is necessary and where the discussion is heading.

Equations

Equations should be treated grammatically as parts of sentences:

The energy E and momentum p of a particle of mass m are related by

$$E^2 = p^2c^2 + m^2c^4 \,, \qquad (2.3)$$

where c is the speed of light. Using Eq. (2.3) one can determine...

Chapter, Section, Equation, etc. uppercase when used with counter.

Mathematical variables are in italics (in LaTeX, use math mode, i.e., \$E^2 = p^2 c^2 + m^2 c^4\$).

Functions such as sin, cos, exp and units are not italicised (in LaTeX, use \sin, \cos, \exp {\rm cm} or \mbox{cm}, etc.).

Subscripts are italicised if they refer to a mathematical variable, e.g., p_x (x is a variable), but not italicised otherwise, e.g., $G_{\rm F}$ (F stands for Fermi, not a variable). In LaTeX: \$G_{\rm F}\$

I, we, one, etc.

I think first person singular sounds awkward.

First person plural can be less awkward but usually not appropriate for a report (who are "we"?).

Best to use a mixture of:

Passive voice: From these results it may be concluded that...

"One": From these results one may conclude...

Or name the subject as appropriate: From these results, the authors of Ref. [37] concluded that...

Or rearrange the sentence, e.g., "...The results of the study led to the conclusion that..."

References

Numbered style: "...the authors of Ref. [17] have concluded..."

In the bibliography, numbered in order of first appearance,

[17] A. Smith and B. Jones, *Investigation of XYZ*, Journal of Interesting Things, 53 (2011) 373-379; e-print arXiv:1107.12345.

Author-year style: "... a recent study (Smith et al., 2011) concluded that..."

In the bibliography references ordered alphabetically by author:

A. Smith and B. Jones, *Investigation of XYZ*, Journal of Interesting Things, 53 (2011) 373-379; e-print arXiv:1107.12345.

To cite a website (e.g., for background info),

[23] Website of the ATLAS Collaboration **`atlas.web.cern.ch`**.

References (continued)

In LaTeX, "... in Ref.~\cite{smith2011}..."

Then in the bibliography (after \begin{thebibliography}):

\bibitem{smith2011} A. Smith and B. Jones, {\it Investigation of XYZ}, Journal of Interesting Things, 53 (2011) 373-379; e-print {\tt arXiv:1107.12345}.

Other useful tools: BibTeX (see e.g. bibtex.org)

If you use a figure from a published source, cite it in the figure caption, e.g.,

Figure 3.3: Invariant mass distribution of photon pairs showing the existence of the Higgs boson (from Ref. [37]).

Always refer to figures/tables in the text itself.

Captions

Caption font should be smaller than text font.

Table 1: Galileo's data on horizontal distance before impact d for five values of the starting height h. The units are punti (points); one punto is slightly less than 1 mm.

h	d
1000	1500
828	1340
800	1328
600	1172
300	800

We will assume the heights h are known with negligible error, and that the horizontal distances d have uncertainties of $\sigma = 15$ punti (points); one punto is slightly less than 1 mm.

Figure caption goes below the figure; table caption above the table.

Code in the report

In the report you should emphasise the mathematics, not the code, e.g.,

> The values of the estimators were determined by minimising the function $\chi^2(\boldsymbol{\theta})$, defined in Eq. (7), with respect to the parameters.

You can include short blocks of code to illustrate specific points as done in the script, e.g., (here with LaTeX verbatim environment)

> The fit function was implemented in python using

```
\begin{verbatim}
def func(x, *theta):
    theta0, theta1 = theta
    return theta0 + theta1*x
\end{verbatim}
```

Miscellaneous

Use British spelling.

Do not cut/paste. We can tell.

Use language precisely. Think carefully about what the words mean:

This derivation shows...

This calculation shows...

This theorem shows...

This equation shows...

Our goal is to get the marked reports with feedback returned to you before you need to write your next report at the end of this module.