Uncertainty notation guidelines for scientific publications.

Greg Henning[1,⋆]

[1] *Université de Strasbourg, CNRS, IPHC UMR 7178, F-67000 Strasbourg, France.*

Abstract. In the following document, I present the motivation for defining a few guidelines regarding uncertainty notation. The guidelines are given in section "The rules"

Why the need for guidelines?

As it turns out, the scientific world has not given itself a set of rules for expressing uncertainties. This is understandable as we cannot reasonably expect all (or even a large number) of fields and teams to have the same requirements, and therefore agree on a fixed set of convention.

Nethertheless, the number of documented conventions is rather small, even when looking at specific places. A quick web search for `uncertainty notation guidelines` yields a surprisingly poor amount of documents going much further than statements like "*The uncertainty of a value is its estimated standard deviation.*"

More surprising, scientific journals such as EPJ's, Physical Review's and even Nature, in their writing guidelines (see in references at the end) barely mention uncertainty. When they do, it is mostly to remind authors that some uncertainties *should* (not "must") be given. Even Wikipedia is short on the subject.

Consequently, the notation and meaning of uncertainties sometimes seems to be an unwritten tradition, passed down from Adviser to Student, generations after generations. **Somewhat fittingly, "uncertainty" often is a confused notion**.

In fact, the reference document on the matter, the **Guide to the expression of uncertainty in measurement** (known as "GUM") gives, in its section 7.2.2, four ways of stating a value with uncertainty:

- x with an uncertainty of u_x,
- x (u_x^{order}), where u_x^{order} is the value of the uncertainty referred to the corresponding last digits of x.
- x (u_x),
- $x \pm u_x$

⋆e-mail: ghenning@iphc.cnrs.fr

Apart from a comment discouraging the last form, the GUM only calls for an explicit and full documentation of the value, its uncertainty and how it was obtained.

This leaves the scientific community with plenty of liberty to write (within the scope of accepted *but unwritten* conventions) the uncertainty in a way he prefers. Unfortunately, the convention used is rarely explicited. The consequences are:

- Given their fuzzy definition, uncertainties are often secondary, both when conducting the work and reading the reading publication.
- Comparing data with given uncertainty can require an important work to understand what the stated numbers represent and put them to the same standard.

Recognizing the fact that any publication needs an explicit and clear set of rules to write and read uncertainty, I propose in the following *a* convention. It is obviously not the the only way (and probably not even the best).

I strongly encourage anyone to sit and consider adopting a standard on uncertainty notation that fits his needs. The following conventions are not fixed and will be updated as needed and according to feedbacks to this paper.

Notation convention

In the following convention, I will follow the principle that *Explicit is better than implicit*

Types of uncertainty

In this convention, I consider two general distribution shape that each deserves its own specific notation: Flat and Gaussian.

- A **flat** uncertainty means the probability distribution of the variable is uniform from a lower to an upper limit around a value (not necessarily symmetric).
- A **Gaussian** uncertainty means that the probability distribution follows a normalized Gaussian shape around the central value (not necessarily symmetric).

Other shapes (triangular, log-normal, Lorenztian, Maxwell–Boltzmann, Poisson, . . .) should be explicited when needed with all the parameters of the distribution

The Rules

1. **Uncertainty is given in full.** *That means that the $x\ (u_x^{order})$ notation is not accepted.* There is no reason why the uncertainty should be given in a shorten way, or that the reader would have to do his own calculations to recover the actual uncertainty number.

2. **Uncertainty is given in the same unit as the value it refers too.** *That means an uncertainty should **never** been given as relative. Relative uncertainty is a comparison tool, not a way of expression.*

3. **Gaussian uncertainty is expressed like** $x\ (u_x)$ *– or* $x({}_{-u_x}{}^{+u_x})$ *in case of asymetric distribution.*
   In $x\ (u_x)$, u_x is the standard deviation (1σ) of the distribution.

4. **Flat uncertainty is expressed like** *$x \pm \delta_x$ – or $x\ {}_{-\delta_x}^{\ \ +\delta_x}$ in case of asymetric distribution.*

5. **Uncertainties are combined by convolution of probability distribution functions** *The final combined uncertainty is not (necessarily) the quadratic sum of its components.*

6. **"*Statistical*" uncertainty refers to an uncertainty from lack of static, but does not presuppose a given shape or value** *It's not because a value is labelled as statistical that it will have a Gaussian uncertainty with a standard deviation of $\sqrt{N}$.*

References

Reference documents on uncertainties:

- GUM: Guide to the Expression of Uncertainty in Measurement
- also available from the Joint Committee for Guides in Metrology
- ICSBEP Guide to the Expression of Uncertainties

Writing guidelines of some scientific publications:

- Style and Notation Guide for Physical Review/Physical Review Letters
- Writing for a Nature journal
- Practical guidance on science journalism and communication
- Nature Physics' Elements of style
- EPJ A's Instructions for authors
- Nuclear Instruments and Methods in Physics Research's Guide for authors