What should be Adopted for Uncertainty of Weighted Average?

Jun Chen

25th NSDD Meeting, 15-19 April 2024

IAEA Headquarters, Vienna

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics under Award Number DE-SC0016948.

Statistical Uncertainty and Systematic Uncertainty

Definitions are taken from "Uncertainty as Applied to Measurements and Calculations" (2011) by J. Denker

Statistical Uncertainty

Uncertainty arising from any measurement of a continuous variable. The true value is not only unknown, but unknowable due to its infinite number of digits. Moreover, there is intrinsic randomness from measurement to measurement.

Can uncertainty in existing data be reduced? ✓

Statistical analysis may be used to get a reduce statistical uncertainty in final result, like averaging results of multiple measurements, or different measurements of the same quantity

---

Systematic Uncertainty

A consistent difference between the indicated and true values, usually arising from a miscalibrated instrument or neglected effect. A systematic uncertainty is always in the same direction as opposed to the random bouncing around characteristic of statistical uncertainties.

Can uncertainty in existing data be reduced? ✗

No systematic uncertainty can be reduced using any statistical analysis, like averaging

More Definitions on Systematic Uncertainty

Systematic effects is a general category which includes effects such as background, selection bias, scanning efficiency, energy resolution, angle resolution, variation of counter efficiency with beam position and energy, dead time, etc. The uncertainty in the estimation of such a systematic effect is called a systematic error.

--- J. Orear, Notes on Statistics for Physicists, UCRL-8417, http://nedwww.ipac.caltech.edu/level5/Sept01/Orear/frames.html

Systematic Uncertainty (error) is reproducible inaccuracy introduced by faulty equipment, calibration or technique.

--- R. Bevington, Data reduction and Analysis for the Physical Sciences. McGraw Hill 1969

Can such uncertainties in existing data, like calibration uncertainties, be reduced after averaging?

Effect of Averaging on Final Uncertainties

Now, do we have a consensus on following statements?

Statement 1:
Averaging as a statistical analysis method may result in a reduced statistical uncertainty for the average value if the input values are in agreement within their uncertainties.

It might be understood this way: better agreement and more statistics give smaller statistical uncertainty

Statement 2:
Averaging as a statistical analysis method cannot result in a reduced systematic uncertainty for the average value no matter how good the input values are in agreement within their uncertainties.

If this is not true, can we accept: better agreement and more statistics give smaller calibration uncertainty? No, it absolutely makes no sense.

Explanations of Weighted Average

Weighted averaging can reduce statistical uncertainty but not systematic uncertainty

(By "reduce", I mean the uncertainty of final result is smaller uncertainty than any input one)

Formula of weighted average in ***statistics theory*** (also in current ENSDF policy document):

For n independent measurements: $x_1 \pm \sigma_1, x_2 \pm \sigma_2, \cdots, x_n \pm \sigma_n$

weighted average $$\bar{x} = \frac{\sum w_i x_i}{\sum w_i}$$

where $w_i = 1/\sigma_i^2$ is the weight of input value x_i

σ_i is the standard deviation (uncertainty) of input value x_i

final uncertainty=MAX(internal unc.,external unc.)

Statistically, the uncertainty in this formula is ***statistical*** by definition

But in real world, the uncertainty of a data value reported in a paper is usually the combination of statistical and systematical uncertainties., which are not given explicitly and separately.

Practical Scenarios for Weighted Average: scenario 1

Scenario 1 (ideal scenario):

Statistical and Systematic uncertainties are given separately for each data value

Step#1: get the weighted average using statistical uncertainty for weight

Step#2: add the final systematic uncertainty to the final statistical uncertainty in Step#1.

Now the question is: how to determine and adopt the final systematic uncertainty, since input systematic uncertainties are usually not the same if they are from different measurements?

Solution#1: take the smallest input systematic uncertainty

Solution#2: take the mean of all input systematic uncertainties

Solution#3: take the largest input systematic uncertainty
(most conservative solution)

But statistical and systematical uncertainties are barely reported separately.

Practical Scenarios for Weighted Average: scenario 2

Scenario 2:

Stat. and Syst. uncertainties are NOT given separately for >=1 data values

No separate final statistical and systematical uncertainties can be obtained

Treat each input uncertainty as it is statistical uncertainty

Then use the average formula to find the final average and uncertainty

This is the most common scenario for performing weighted average by researchers and evaluators.

But the problem is that, it naturally assumes that all input uncertainties are treated as statistical uncertainties and could give an overall reduced final uncertainty, while **in many cases systematical uncertainties are significant or even dominant and they can't be reduced by averaging**.

Weighted averaging can reduce statistical uncertainty but not systematic uncertainty

So in this scenario, the final uncertainty could be underestimated for average of data values in agreement.

Practical Scenarios for Weighted Average: scenario 3

Scenario 3:

Stat. and Syst. uncertainties are NOT given separately for >=1 data values
No separate final statistical and systematical uncertainties can be obtained
} same as scenario 2

Treat each input uncertainty as it is statistical uncertainty, BUT only for the purpose of using the average formula (for getting the average value)

Then adopt the final uncertainty as:

final uncertainty=MAX(internal unc.,external unc., smallest input unc.)

The difference of uncertainty between Scenario 2 and 3 is that:

Scenario 2<=actual uncertainty<=Scenario 3

Scenario 3 could give a larger uncertainty (smallest input unc.) than the former (Scenario 2) which gives a reduced "statistical" uncertainty as the final uncertainty for data values in agreement.

It is safer to adopt final uncertainty in Scenario 3

Currently, it is only recommended explicitly for averaging half-life values in the ENSDF guideline.

Summary and Conclusions

Weighted averaging can reduce statistical uncertainty but not systematic uncertainty

Final uncertainty in different scenarios:
Scenario 2<=actual uncertainty<=Scenario 3

S2: unc=MAX(ext,int)
S3: MAX(S2 unc, smallest input unc)

Scenario 3 approach is recommended explicitly for averaging half-life values in the ENSDF guideline.

final uncertainty=MAX(internal unc.,external unc., smallest input unc.)

Current situation for adopting the final uncertainty:

both Scenario 2 and Scenario 3 approaches are used by different evaluators resulting in a significant inconsistency in ENSDF evaluation.

Current policy document only gives the formula of average, but not address any of those scenarios and problems, leaving ambiguity in adopting uncertainty of weighted average.

A more specific, clear and consistent policy is needed for adopting uncertainty of weighted average