Tutorial on error treatment

Giorgio Levy

UBC

September 22, 2014

Uncertainty of a measurement

Reporting the result of a measurement

The result of a measurement is reported by the central value, the error, and the magnitude, *i.e.* 1.0 ± 0.4 mm.
To round the reported results we may follow the guidelines given by the Particle data group:

- If the three highest order digits of the error lie between 100 and 354, we round to two significant digits.
- If they lie between 355 and 949, we round to one significant digit.
- If they lie between 950 and 999, we round up to 1000 and keep two significant digits.

In all cases, the central value is given with a precision that matches that of the error. For example, the result (coming from an average) 0.827 ± 0.119 would appear as 0.83 ± 0.12, while 0.827 ± 0.35 would turn into 0.8 ± 0.4.

Resolution & Sensitivity

Resolution

The resolution is the smallest change on the measured quantity that can be detected by the sensor.

Sensitivity

The sensitivity is the smallest magnitude of change in the quantity being measured that can be detected in the sensor. It is associated to response (at a given resolution) of the apparatus.

Precision & Accuracy

Precision

The degree to which repeated measurements under unchanged conditions show the same results.

Accuracy

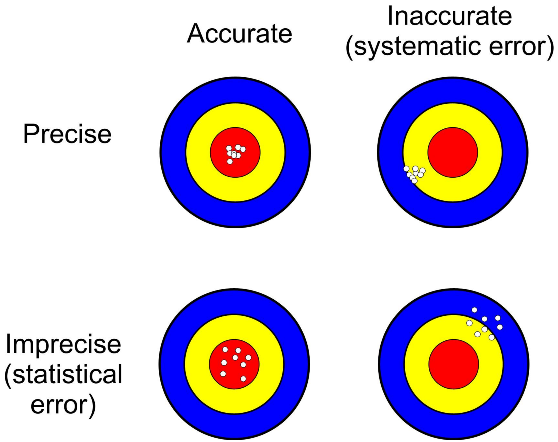
The degree to which the result of a measurement, calculation, or specification conforms to the correct value or a standard.

Precision & Accuracy

To improve the level of precision and accuracy, the experimentalist should aim to reduce both the statistical and the systematical uncertainty.

- Repeating measurements reduces the statistical uncertainty and leads to precise measurements (small spread of values).
- Calibrating the equipment reduces the systematic uncertainty and leads to accurate measurements (central value close to the true value).

Precision & Accuracy



Random errors

Central limit theorem

The distribution of the mean $\bar{x}$ with variance σ^2 of a set x_i of N independent and identically distributed variables tends to a normal or Gaussian distribution $G(x)$, regardless of the underlying distribution for x_i.

$$G(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$$

Law of large numbers

The average of the results from a random process converges to the distribution mean as the number of trials increases.

Standard deviation & Standard error

Mean

The mean of N repeatable measurements of a quantity x is:

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

Standard deviation

The standard deviation σ is

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(\bar{x} - x_i)^2}$$

Standard error of the mean

The standard error of the mean δx– or simply the standard error – is

$$\delta x = \frac{\sigma}{\sqrt{N}}$$

Standard deviation & Standard error

SD

The standard deviation is a property of the method and the equipment. It does not decrease with increasing number N of repeatable measurements.

SE

The standard error defines the precision on the mean. It decreases with the number N of repeatable measurements.

Error propagation

Derivation for the propagation formula

For variables x and y with errors δx and δy, the error δz of a variable z, which is a function of x and y: $z = f(x, y)$, the value at $(\bar{x}, \bar{y})$ is

$$\bar{z} = \frac{1}{N}\sum_{i=1}^{N} f(x_i, y_i).$$

We define the error in z, δz, as

$$(\delta z)^2 = \frac{1}{N(N-1)}\sum_{i=1}^{N}(z_i - \bar{z})^2,$$

$$(\delta z)^2 = (\partial_x f \delta x + \partial_y f \delta y)^2$$

Where we have expanded $f(x, y)$ in a Taylor's series around $(\bar{x}, \bar{y})$:

$$\begin{aligned} f(x,y) &\simeq f(\bar{x},\bar{y}) + \partial_x f|_{(\bar{x},\bar{y})}(x-\bar{x})\ldots \\ &+ \partial_y f|_{(\bar{x},\bar{y})}(y-\bar{y}) \end{aligned}$$

Which gives

$$\begin{aligned} (\delta z)^2 &= (\partial_x f)^2(\delta x)^2 \ldots \\ &+ (\partial_y f)^2(\delta y)^2 \ldots \\ &+ 2\partial_x f \partial_y f \delta(x,y), \end{aligned}$$

where we define the covariance $\delta(x, y)$ as

$$\delta(x,y) = \frac{1}{N(N-1)}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})$$

Error propagation

General formula

For variables x and y with errors δx and δy, the error δz of a variable z, which is a function of x and y: $z = f(x, y)$, is calculated as:

$$(\delta z)^2 = \left(\left.\frac{\partial f}{\partial x}\right|_{\bar{x},\bar{y}} \right)^2 (\delta x)^2 + \left(\left.\frac{\partial f}{\partial y}\right|_{\bar{x},\bar{y}} \right)^2 (\delta y)^2 + 2 \left.\frac{\partial f}{\partial x}\right|_{\bar{x},\bar{y}} \left.\frac{\partial f}{\partial y}\right|_{\bar{x},\bar{y}} \delta(x, y).$$

This formula can be used to deduce the following rules:

Addition or subtraction

The absolute errors are added in quadrature.

$z = x + y \longrightarrow (\delta z)^2 = (\delta x)^2 + (\delta y)^2$

Multiplication or division

The relative errors are added in quadrature.

$z = xy \longrightarrow \left(\frac{\delta z}{z}\right)^2 = \left(\frac{\delta x}{x}\right)^2 + \left(\frac{\delta y}{y}\right)^2$

Weighted mean

The weighted mean of x_i results with error δx_i is

$$\bar{x} = \frac{\sum_{i=1}^{N} \frac{x_i}{(\delta x_i)^2}}{\sum_{i=1}^{N} \frac{1}{(\delta x_i)^2}}$$

and the error on the weighted mean $\delta \bar{x}$ is

$$\delta \bar{x} = \sqrt{\frac{1}{\sum_{i=1}^{N} \frac{1}{(\delta x_i)^2}}}$$

Least-squares fit

Least-squares fit

Given a set (x_i, y_i) measurements where y is a dependent variable of x through the function: $y = f(p_j, x)$, we would like to determine the parameters p_j that minimize

$$\chi^2 = \sum_{i=1}^{N} \frac{[y_i - f(p_j, x_i)]^2}{(\delta y_i)^2}.$$

The most used method is the LevenbergMarquardt algorithm however other techniques can be used like genetic algorithms, simulated annealing. We only focus on the first method. The uncertainty on the parameters is calculated through the diagonal term of the covariance matrix. It is normally given by the package that implements the algorithm.

References

References

- Data reduction and error analysis for the physical sciences, by P.R. Bevington and D.K. Robinson, McGraw-Hill Science Publications.
- Statistical data analysis, by Glen Cowan, Oxford Science Publications.