Introduction to Error Analysis

Ajay K. Prasad
Associate Professor
Department of Mechanical Engineering
University of Delaware

Introduction

Error analysis is a critical aspect of experimentation. Most technical journals nowadays will not accept experimental results unless the data on a plot are accompanied by error bars. Unfortunately, error analysis is not always given the respect and appreciation it deserves, and is treated by some as a nuisance. Placing error bounds on a result is sometimes difficult, but it conveys a wealth of information to the reader. Therefore, in this series of experiments, we will give significant weight to error analysis.

To begin with, let us review some concepts, **accuracy** and **precision**. Accuracy represents how close the *measured* value is compared to the *true* value. Precision represents the amount of *scatter* in the data. For example, consider an experiment in which a darts are thrown one by one at a dart board (Figure 1). If all the darts end up in a tight bunch, but do not land on the bulls-eye, then this would be an inaccurate but precise measurement. If instead, they are centered on the bulls-eye, but show significant scatter, then they are accurate but imprecise. Only when they are tightly bunched together and centered on the bulls-eye, can we say that the measurement is both accurate and precise.

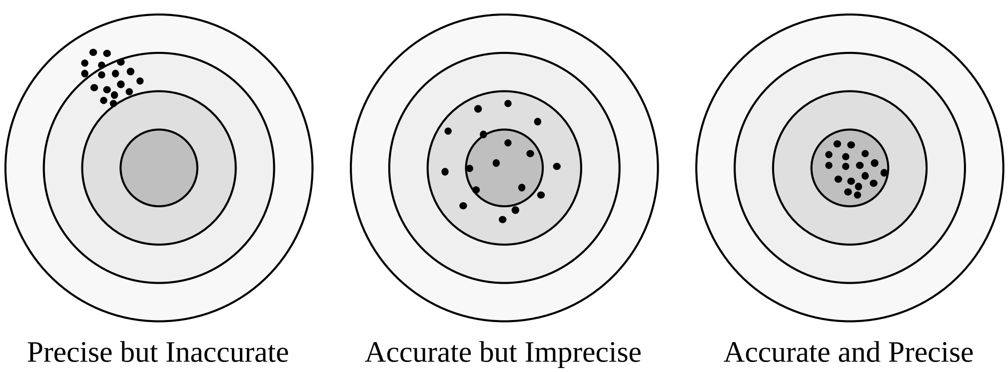

Figure 1: Illustration of precision and accuracy

Systematic and Random Errors

Errors are classified into **systematic** errors and **random errors**. Systematic errors lead to inaccurate results. They cause a bias or a shift in the results away from the true value because they always overestimate (or always underestimate) the quantity being measured. For example, if one were to make several, very careful measurements of the period of oscillation of a pendulum, but with a stopwatch that runs fast, one would get an erroneous result (precise but not accurate). Some features of systematic errors:

1. Systematic error is reproducible, and hence not easy to detect.
2. It cannot be reduced by taking more measurements.
3. Can be uncovered only by careful calibration of the instrument.

Random error results in scatter in the data. For example if one were to repeat the pendulum period experiment with a well-calibrated, high-quality stopwatch, but after consuming several cups of coffee, the resulting "trigger-happiness" will result in higher random error (results will not be precise). Some features of random error:

1. Varies in magnitude; equally likely to be positive or negative.
2. It is not reproducible on a one to one basis.
3. It is very apparent.
4. Can be reduced by taking more measurements. However, the reduction process is time consuming, as the statistical uncertainty decreases as $1/\sqrt{\text{time}}$. Can also be reduced by improving measurement technique (avoid the coffee).

In summary, it is easy to see that:

systematic errors lead to inaccurate results,
whereas,
random errors lead to imprecise results

Mean and Standard Deviation of a Sample

Consider a set of N measurements x_i with $i = 1, 2, \cdots, N$. The *sample mean* $\overline{x}$ is defined by

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

As N becomes large, $\overline{x}$ approaches the true value (provided systematic errors are ignored).

The *standard deviation*, σ_x is given by

$$\sigma_x^2 = \frac{1}{N} \sum_{i=1}^{N} (x_i - \overline{x})^2 = \overline{x^2} - (\overline{x})^2 \tag{2}$$

Note that σ_x^2 is also referred to as the *variance*.

After N measurements are made, one can determine $\overline{x}$ and σ_x. If another measurement is now made, how far from the mean value will it fall? σ_x determines the range into which this new measurement is likely to fall. The new measurement has a 67% probability of lying in the range $\overline{x} \pm \sigma_x$, a 96% probability of lying in the range $\overline{x} \pm 2\sigma_x$, and a 99.7% probability of lying in the range $\overline{x} \pm 3\sigma_x$, provided N is large.

Standard Error or Standard Deviation of the Mean

$$\sigma_{\overline{x}} = \frac{\sigma_x}{\sqrt{N}} \tag{3}$$

While σ_x is the standard deviation of a single measurement, x_i, $\sigma_{\overline{x}}$ is the standard deviation of the mean value of N measurements, $\overline{x}$. $\sigma_{\overline{x}}$ is obviously smaller than σ_x because the mean value of N measurements is more reliable than a single measurement.

Correlation coefficient r

In several experiments, you are asked to compare the measured value m_i with the corresponding theoretical value t_i with $i = 1, 2, \cdots, N$ where N is the number of data points. The values of t_i plotted against the values of m_i should fall on a 45° line if the theory is completely correct and there are no experimental errors. To quantify the degree of the agreement between m_i and t_i, we usually use the correlation coefficient r between m_i and t_i defined as

$$r = \frac{\sigma_{mt}}{\sigma_m \, \sigma_t} \tag{4}$$

with

Variance of m_i:

$$\sigma_m{}^2 = \frac{1}{N}\sum_{i=1}^{N}(m_i - \overline{m})^2 = \overline{m^2} - (\overline{m})^2 \tag{5}$$

Variance of t_i:

$$\sigma_t{}^2 = \frac{1}{N}\sum_{i=1}^{N}(t_i - \overline{t})^2 = \overline{t^2} - (\overline{t})^2 \tag{6}$$

Covariance:

$$\sigma_{mt} = \frac{1}{N}\sum_{i=1}^{N}(m_i - \overline{m})(t_i - \overline{t}) = \overline{mt} - \overline{m}\,\overline{t} \tag{7}$$

where $\overline{m}$ and $\overline{t}$ are the mean values of m_i and t_i, respectively.

The perfect agreement between m_i and t_i corresponds to $r = 1$, while $r = 0$ indicates no agreement whatsoever. It is noted that the correlation coefficient can also be used to estimate the correlation between two statistical variables. On the other hand, the correlation coefficient is needed herein to evaluate the degree of agreement between the theory and experiment.

A package such as MS Excel will calculate the correlation coefficient for you directly. Please see the relevant program function description for details.

t-Test

The ongoing discussion actually pertains to the case when the sample size is rather large. In practice, one does not often have the time to perform such a large set of measurements. When the sample size is smaller, it turns out that the error bar becomes larger for the same confidence limit. You may refer to the t-table in any mathematical or statistical handbook for the relation between sample size, error range and the corresponding confidence limit.

For example, for a large sample, a 95% confidence limit corresponds to $\pm 1.96\sigma$. If the sample size were smaller, say 30, 20, or 10, then the error range would have to be increased to $\pm 2.04\sigma$, $\pm 2.09\sigma$, or $\pm 2.23\sigma$, respectively for the *same* confidence limit of 95%.

Propagation of errors

Most often, it is not possible to directly measure the desired quantity, but it must be obtained by taking products, sums, ratios, etc., of a number of measured quantities. For example, consider the case where you measure flow rate by collecting X liters of water in Y seconds. The desired quantity M is X/Y. What is the error ΔM in M, given the error ΔX and ΔY in X and Y respectively?

First of all, ΔX and ΔY can be estimated by examining the instrument. For example, if the *least count* (smallest division) of a ruler is 1 mm, then the measurement error is less than 1 mm, usually 0.5 mm. If you use a magnifying lens to record the measurement very carefully, you can probably reduce it to even 0.2 mm.

However, if you use a stopwatch, it is not sufficient to use the least count of the stopwatch as your error estimate. Why? What about the human reaction time involved in starting and stopping? A good error analysis will carefully assign values to these quantities.

To find the propagated error, one must know whether the errors ΔX and ΔY are systematic or random in nature (or a mixture of the two), and also the functional relationship between M and X and Y.

Propagation of Systematic Errors

There are three types of systematic errors:

- of the same sense
- of unknown sense
- of opposite sense

If X and Y are the same physical quantity, and if they are measured with the same instrument, then the systematic error will make both measurements either too large or too small. Therefore, they will be of the same sense. If X and Y are different types of quantities and measured with different instruments, then the errors will be of unknown sense. The case where X and Y are of opposite sense is usually rare in practice. (Why? If you know a measurement is too large, why not compensate for it?)

Consider the following functional dependencies:

1. $M = X + Y$

   The error is $(\Delta M)_{sys} = \Delta X \pm \Delta Y$.

   The relative error $\Delta M/M$ is given simply as:

$$\left(\frac{\Delta M}{M}\right)_{sys} = \frac{\Delta X \pm \Delta Y}{X+Y} = \frac{X}{X+Y}\left(\frac{\Delta X}{X}\right) \pm \frac{Y}{X+Y}\left(\frac{\Delta Y}{Y}\right)$$

   - If the errors are of the same sense, choose the positive sign.
   - If the errors are of opposite sense, choose the negative sign.
   - If the errors are of unknown sense, the worst case scenario must be chosen, which is the same as choosing the positive sign.

2. $M = X - Y$

   The error is $(\Delta M)_{sys} = \Delta X \mp \Delta Y$.

   The relative error $\Delta M/M$ is given simply as:

$$\left(\frac{\Delta M}{M}\right)_{sys} = \frac{\Delta X \mp \Delta Y}{X-Y} = \frac{X}{X-Y}\left(\frac{\Delta X}{X}\right) \mp \frac{Y}{X-Y}\left(\frac{\Delta Y}{Y}\right)$$

   - If the errors are of the same sense, choose the negative sign.
   - If the errors are of opposite sense, choose the positive sign.
   - If the errors are of unknown sense, the worst case scenario must be chosen, which is the same as choosing the positive sign.

   Note that if X and Y are about the same magnitude, the relative error can become huge! This simple example demonstrates how difficult it is to measure the difference in large quantities of equal magnitude.

3. $M = X/Y$

   Now,

$$(\Delta M)_{sys} = \frac{1}{Y}\Delta X - X\frac{\Delta Y}{Y^2}$$

   Therefore,

$$\left(\frac{\Delta M}{M}\right)_{sys} = \frac{\Delta X}{X} - \frac{\Delta Y}{Y}$$

   $M = X/Y$ is a typical relative-size measurement. If X and Y are measured with the same instruments, then the fractional error in M is small, even if the fractional error in X and Y is large.

4. $M = XY$

$$\left(\frac{\Delta M}{M}\right)_{sys} = \frac{\Delta X}{X} + \frac{\Delta Y}{Y}$$

5. $M = aX^b$

   Here, a and b are given constants. Then,

$$\left(\frac{\Delta M}{M}\right)_{sys} = b\frac{\Delta X}{X}$$

Propagation of Statistical Errors

Statistical (or random) errors can be of either sign. Hence when these errors are added, divided, multiplied, etc., they must be treated as errors of unknown sense to each other. They *always* combine to make the final uncertainty larger. The resultant error is the square root of the sum of the square of the error components (RSS: Root of the Sum of the Squares).

Let the functional dependence between M the final desired quantity, and the individual contributing measurements be

$$M = f(X, Y, Z, \cdots)$$

Then,

$$(\Delta M)_{st} = \sqrt{\left(\frac{\partial f}{\partial X}\Delta X\right)^2 + \left(\frac{\partial f}{\partial Y}\Delta Y\right)^2 + \left(\frac{\partial f}{\partial Z}\Delta Z\right)^2 + \cdots}$$

You should be able to show these for yourself:

- $M = X \pm Y$

$$\left(\frac{\Delta M}{M}\right)_{st} = \sqrt{\left(\frac{X}{X \pm Y}\right)^2 \left(\frac{\Delta X}{X}\right)^2 + \left(\frac{Y}{X \pm Y}\right)^2 \left(\frac{\Delta Y}{Y}\right)^2}$$

- $M = XY$ or $M = X/Y$

$$\left(\frac{\Delta M}{M}\right)_{st} = \sqrt{\left(\frac{\Delta X}{X}\right)^2 + \left(\frac{\Delta Y}{Y}\right)^2}$$

When Both Systematic and Statistical Errors are Present

In this case, first determine the errors from each cause separately, and report the final error as the RSS of the individual errors.

Always state the confidence limit of your error. If you say that the final fractional error is 5%, does this correspond to $\sigma_{\overline{x}}$, 2 $\sigma_{\overline{x}}$, or 3 $\sigma_{\overline{x}}$? Typically, in scientific work, we use the 67% confidence limit, i.e., $\sigma_{\overline{x}}$.

The error should be stated to only one significant figure. The measured quantity should be rounded off so that its least significant figure corresponds to the most significant figure of the error. Examples: 98.2 cm ± 0.2 cm, not 98.21 cm ± 0.2 cm (the 0.01 cm is insignificant compared to the error!).

References

John R. Taylor, “An Introduction to Error Analysis.”