EXPERIMENTAL ERRORS

The following is a very simple introduction to the topics of Systematic and Random Error, subjects that are of paramount importance to the physical scientist.

There are primarily two kinds of errors that affect experimental results:

1) Systematic Error

Systematic error refers in a collective sense to a broad range of errors that arise due to imperfect equipment design and/or operator error. By operator error we also include inexact computational procedures used during the data reduction process. We do not, however, include accidental blunders and arithmetic mistakes which are remediable.

Properties of Systematic Errors
Systematic errors are characterized by the following properties:

i) **They are repeatable from experiment to experiment, with respect to both the sign of the error and also its approximate magnitude**
ii) **They are not amenable to any kind of statistical analysis**
iii) **They cannot be reduced or detected by repeating an experiment**

Examples:
1) Defective meter zero in reading an instrument
2) Incorrect positioning of the fiducial mark in the manufacture of a pipet
3) Ignoring the stem correction when using a thermometer or the temperature correction for the metal scale when using a barometer
4) Using overlarge tires on a car thus creating a constant error in the speedometer reading
5) Using an impure sample as a calibration standard (e.g., KHT) when determining the normality of an acid or base
6) Using an out-of-date value for a physical constant such as the gas constant R or the gravitational constant which varies with location
7) Use of an uncalibrated balance
8) Early truncation of an infinite series or failure to take into account computer roundoff error in the evaluation of an infinite series etc.

Each of these examples will cause an error in an experimental result, either directly during the course of the experiment or later during the data reduction procedure. Furthermore, every time the experiment is repeated, these errors will continue to enter, usually to about the same extent and with the same positive or negative sign.

Detection and elimination of systematic error is very difficult because one is rarely certain of its origin or, for that matter, its existence. One might, for example, rely on comparison measurements between different laboratories of the quantity in question, especially if the measurements are intended to be highly accurate. More routinely, the use of blank determinations or the use of certified standards (e.g., the use of a NBS standard Benzoic Acid sample for the calibration of a bomb calorimeter) are obvious strategies to reduce systematic error but no procedure, however extreme, will guarantee an error free measurement. In fact **it is impossible to ever know for certain that all systematic error has been eliminated from an experiment!**

Measures of Systematic Error
There are no formal measures. The best you can do is estimate systematic error based on comparisons against known standards or known instrument shortcomings.

Relationship to Accuracy

SYSTEMATIC ERROR IS THE LIMITING FACTOR IN DETERMINING ACCURACY

Thus systematic error, which is always present to some extent, will ultimately determine the accuracy of a measurement; a concept we now formalize.

Definition of Accuracy
The notion of accuracy is based on the concept of the 'true value' of a measurement, i.e., the quaesitum. We denote this quantity, which is the holy grail of any measurement, as τ. For reasons already discussed, one can never know τ but we can at least hope that the mean of our measurements will approximate it. If we define the difference between the mean μ of an infinite number of measurements and the true value τ, as the bias,

$$\text{bias} = \mu - \tau \quad \text{(assuming an infinite number of measurements)}$$

then we can assert the following: 'when the bias is zero, the measurements have no systematic error and are therefore said to be '***accurate'.*** This simply formalizes the relationship between systematic error and accuracy. Of course we cannot know τ (ever!), so of course we can never really know if the bias is zero and therefore if our results are truly accurate.

2) Random Error

Repeated measurements of some physical quantity will always exhibit fluctuations of the measured values about some mean value. Indeed, it is a fundamental assumption (sometimes called the Postulate of Measurement) that, as the number

of measurements tend to infinity*, the running mean will eventually settle down to some fixed quantity called the 'limiting mean'. If this assumption were not true, it would be impossible to meaningfully assign a value to any variable, and it would certainly vitiate the whole concept of measurement.

Furthermore, the greater the resolution of the measuring apparatus, the greater the extent of these fluctuations. These fluctuations are assumed to arise because of a number of small inherently uncontrollable processes, each of whose effect is individually small. These errors are thus called **Random Errors** and can be reckoned as the difference between the value of the ith measurement and the mean value, i.e., $\varepsilon_i = x_i - \overline{x}$ where ε_i is the ith random error associated with the ith measured value. Note that ε_i can be positive or negative.

Properties of Random Error
1) The random errors associated with a specific measuring apparatus are a characteristic of the design of that particular apparatus.
2) The sign of the individual errors, ε_i, is assumed to be positive or negative with equal probability.
3) These errors, when tabulated for a very large number of measurements, will be assumed to occur with a frequency that follows a normal distribution, that is, if one plots $f(\varepsilon)$ vs. ε, the resulting curve will be approximately Gaussian in nature. This means that random error can be analyzed statistically.
4) Unlike systematic errors, random error is detected by simply observing that the measured values fluctuate about the mean and therefore, by changing the design of the measuring apparatus, they can be reduced. However, they can never be totally eliminated.

Examples of Random Error
1) Noise effects in electrical circuits (always present)
2) Brownian motion
3) Any 'on/off' cyclic event (e.g., an electrical heater) that is part of the measuring apparatus provided that the event behaves randomly relative to the time in which a measurement is made. If the timing of the measurement and the duty cycle of the 'on/off' event are synchronous, the error is more likely to be systematic in nature. This is an example showing that the characterization of what is random and what is systematic is not always clear-cut.

Measures of Random Error
Because random error is amenable to statistical analysis, it is possible to derive a measure of its influence. We distinguish between random error in the individual

---

* Symbolically, $\lim\limits_{n \to \infty} \overline{x}(n) \to \mu$

where μ is usually denoted as the population mean, i.e., the mean value for the infinite experiment.

measurements $x_1, x_2, ..., x_n$ and the random error associated with the mean value $\bar{x}$, computed from the individual values.

The most common measures of the precision in the original data points are the sample VARIANCE and the corresponding STANDARD DEVIATION, viz.,

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 \quad \text{Sample Variance}$$

or

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2} \quad \text{Sample Standard Deviation}$$

where $\bar{x}$ is the sample mean defined by:

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

For the mean itself, the preferred measure of uncertainty is the STANDARD ERROR, defined as: (the $\cong$ denotes the sample estimate)

$$S.E = \sqrt{\frac{\sigma^2}{n}} \cong \sqrt{\frac{S^2}{n}} = \frac{S}{\sqrt{n}} = \sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Notice that repeating experiments reduces the standard error by the factor $1/\sqrt{n}$. Thus while the standard deviation (for the infinite experiment) is a characteristic of a particular apparatus and therefore can be reduced only by redesign of the apparatus, the standard error can be reduced by repetition of experiments. When dealing with the mean of a series of experiments, the standard error is actually the best estimate of random error so experimental repetition is often well worthwhile. On the other hand, the reduction in the standard error is not linear with respect to n so at some point it may be more economical to look at improving the apparatus instead of just making more measurements.

Relationship to Precision

**RANDOM ERROR IS THE LIMITING FACTOR IN DETERMINING PRECISION**

Definition of Precision
Precision reflects how close our measurements are to one another. If we make a series of measurements and then calculate the standard deviation for these results we will find that the smaller the standard deviation the better our notion of precision, whereas conversely, the more spread out the results, the larger the

standard deviation and hence the poorer the precision. The phrase 'precise', like the word 'accurate' is entirely dependent on what the experimentalist is willing to accept and has no quantitative meaning beyond that. Incidentally, the word 'precise' is sometimes used *imprecisely* to denote 'accurate' as well as 'high resolution' (presumably a lot of significant digits).

In summary then, accuracy measures how close the mean value is to the true value whereas precision measures the closeness of the individual results. Repetition will increase precision by reducing the standard error, but will have no effect on accuracy.

The following diagrams illustrate these points. In the case of good precision and poor accuracy, an improperly adjusted sight would explain the systematic error causing a non-zero bias. Once the sight is adjusted, the bias goes to zero. In the case of poor accuracy and poor precision we have both systematic and random error to contend with and it is difficult to know which is which. The systematic error in the horizontal direction could be explained by a strong and variable wind while the vertical error could be explained by defective ammunition. Of course the guy could also be just a lousy shot.
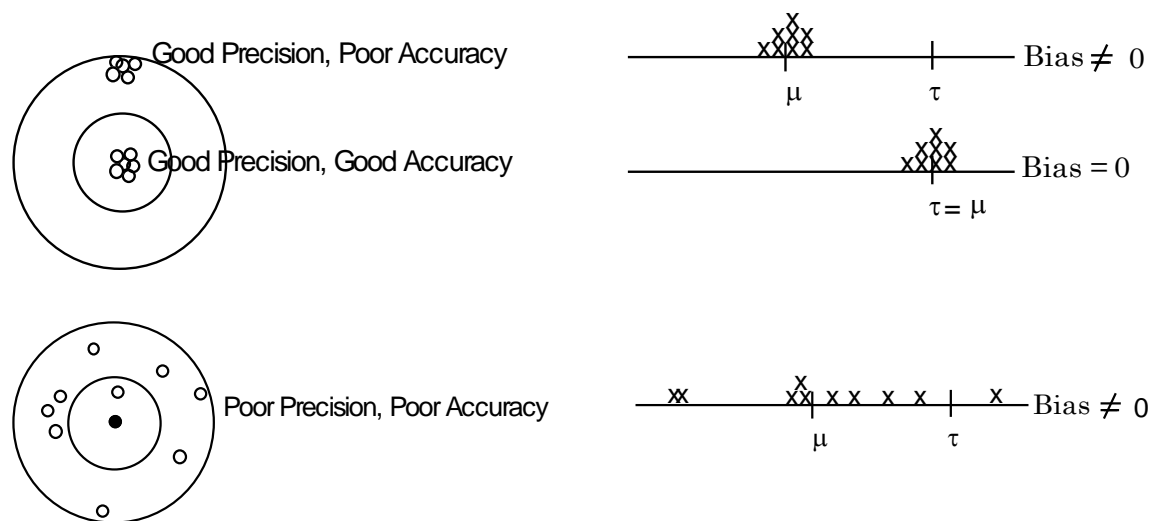
These diagrams also make the point that just because a series of measurements is precise, no conclusion whatsoever can be drawn about the accuracy of the measurements unless one can also be certain that there is little likelihood of systematic error.

Method of Stating Error in a Measurement
As already indicated, the standard error is the appropriate measure of the uncertainty when the mean value is used to express an experimental result. It is recommended that uncertainty be expressed as follows:

$$X = \overline{X} \pm s.e.$$

where the mean and the standard error were defined previously. Furthermore, the standard error, s.e., should be expressed to no more than two significant digits and should agree with $\overline{X}$ as to the number of significant digits after the decimal. It is also important to state the number of measurements, since the standard error depends on n, as well as make some statement as to the experimentalist's belief in the magnitude of the systematic error.

Numerical Example

Suppose we wish to measure the density of a newly synthesized liquid compound at a fixed temperature of 25° C. We obtain the following four values: 1.3271, 1.3303, 1.3100 and 1.3210 where the results are expressed in the units of gms cm^{-3}.

Using the equations given above, we compute the following statistics:

$$\overline{d} = \frac{1}{4}\sum_{i=1}^{4} d_i = 1.3221$$

$$S^2 = \frac{1}{4-1}\sum_{i=1}^{4}(d_i - \overline{d})^2 = 7.9953 \text{ x } 10^{-5}$$

$$S = 8.94166 \text{ x } 10^{-3}$$

$$s.e. = \sqrt{\frac{S^2}{n}} = \frac{S}{\sqrt{n}} = \frac{8.94166 \text{ x } 10^{-3}}{\sqrt{4}} = 0.00447$$

Note that we have carried one extra significant figure when a computed value exceeds five digits. Rounding is done at the end, NOT at each step of the calculation.

We now express our measured result in the form:

$$\text{Measured Density} = \overline{d} \pm s.e. = 1.3221 \pm 0.0045 \quad (gms\ cm^{-3})$$

where the expressed uncertainty in the density is the sample estimate of the standard error computed for a total of four measurements assuming a negligible systematic error.

This last statement is an integral part of the error statement and should not be omitted (but usually is, unfortunately).

How do we know the systematic error is negligible? In fact we don't for reasons already discussed, but it probably is very small IF we have made a conscientious effort to account for each possible error in the experimental process. One technique would be to make a series of measurements using this apparatus on a liquid of

known density, such as water, and then use whatever discrepancy existed between accepted values in the literature and our measurements as an indication of possible systematic error. Even so, systematic error is very elusive.

8-24-05