Measurement Errors

It is common for students of science to think that the most important outcome of an experiment is a measured value. But the sages know that the more important – and difficult to determine – outcome of an experiment is the **uncertainty** of that measured value. The only thing known with absolute certainty is that the measured value is not the true value. However, if the uncertainty is properly determined, the true value will lie within the interval of uncertainty around the measured value. Similarly, it makes no sense to compare the measured values of two different experimental teams – they assuredly will differ. Rather what is to be compared for overlap are the intervals of uncertainty of two experiments.

Put simply, "the wise person knows well how wrong s/he might be".

Random Errors and Systematic Errors

No matter how good your measurement tool and method, at some level there will be **measurement errors**. In scientific jargon these errors are *not* mistakes but rather "unavoidable" variations. Note though that experimental science is primarily a battle to reduce such errors, so that what is unavoidable in one experiment is avoided in the next.

There are two fundamentally different types of measurement errors:

> **Random errors** deviate in either direction from the true value with equal probability. The size of a random error will vary with each measurement, but a histogram of errors from many repeated measurements will have a characteristic distribution called a Gaussian distribution with a characteristic width called the standard deviation σ (Figure 1). The size of random errors (or equivalently the width of the error distribution) determines the **precision** of an experiment. Random errors have the nice property that if you average together many measurements, the average value will be closer to the true value than a single measurement.[1]
>
> **Systematic errors** always deviate in the same direction. Unnoticed, they will lead to incorrect answers no matter how many measurements are made. A classic example is measuring the height of a wall with a "meter" stick that is actually 95 cm long. No matter how carefully you measure, and no matter how many measurements you average together, your result will still be incorrect. The size of systematic errors determines the **accuracy** of an experiment.

Note that an experiment can be very precise *and* very inaccurate! That is, the range of random errors may be very small but still not include the true value.

[1] Just to exercise your understanding, the more formal way to put this is that the standard deviation of the mean (or average) is smaller than the standard deviation of each measurement.

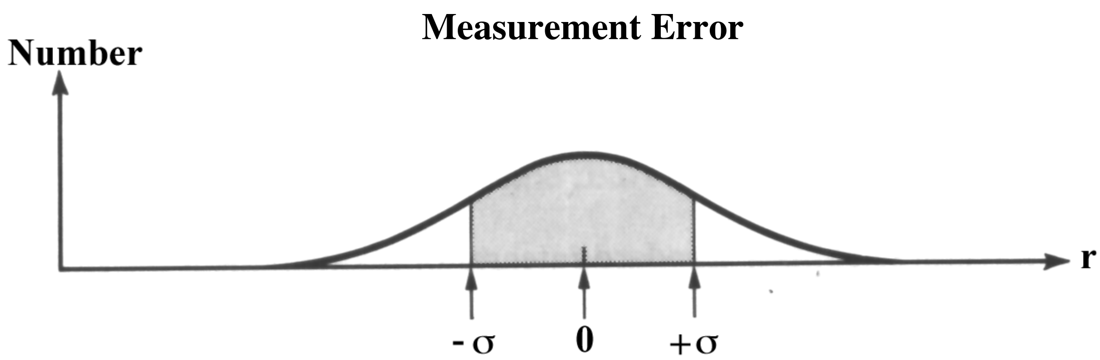

Figure 1. Gaussian distribution

Error Estimation

If multiple measurements can be made, making an estimate of random errors is straightforward. Essentially all that needs to be done is to examine the variation of the several measurements. To obtain a very rough estimate of how far off the measured value might be from the true value, one might use half the range of the measurements. A more sophisticated (but not difficult) approach is to compute the standard deviation of the measurements.

Systematic errors are far more insidious, for at the scientific forefront the true value is not known. Only confirmation by a diverse set of experimental teams and methods can provide some assurance that a result is not corrupted by systematic errors.

Typically, the random error for a measurement is quoted as one standard deviation. But it is very important to appreciate that random errors are distributed according to a Gaussian distribution and do not have a single value. The standard deviation is just a "typical" deviation of a measurement from the mean. Some measurements will deviate more, some less. **For a Gaussian distribution, two thirds of your measurements will be within one standard deviation of the mean** (Figure 1). Turning this fact around, each time you make a measurement there is a 1 in 3 chance that the measurement is in error by more than one standard deviation. There is a 1 in 20 chance that the measurement is in error by more than two standard deviations. There is a 1 in 370 chance that the measurement is in error by more than three standard deviations.

Consider the implications of this deeply, for this is the essence of risk analysis! However careful one may be, there is a 1 in 20 chance that a solitary measurement will be in error by two standard deviations. Suppose such an error is unacceptable – it may lead to a bridge failure or a plane crash. Then the precision of your experiment must be improved. But better precision costs money and time. How large a margin of error are you willing to live with? And consider further … suppose the determination of the measurement is optimistic. Then an error thought to be two standard deviations may in fact be only one standard deviation, with a 1 out of 3 chance of happening. Again, "the wise person knows well how wrong s/he might be".

Better Precision through Averaging

It is a basic practice to improve precision by making multiple measurements and then taking the average – or **mean** - for the final measured value. *If* the variations in your measurements are random, the more measurements you make the more reliable should be your computed mean. For example, imagine computing a mean from two measurements, one of which just happens to be the very deviant from the true value. (Of course, you can't know this, since you don't know the true value.) Clearly the mean will be better than the single deviant measurement. Now imagine computing the mean from those two measurements plus three more. Even though both computations include the most deviant measurement, the mean from all five measurements will be closer still to the true answer. Colloquially, this is referred to as "averaging out" errors.[2]

Final Thought

While you may never have to compute a standard deviation in "real" life, measurement is something that we all do. Maybe its for building a loft, or timing an event, or weighing produce. Whatever, using the concepts of random errors, systematic errors, and computing means from multiple measurements can improve the precision and success of your efforts.

The wise person knows well how wrong s/he might be

[2] Of course the mean also will not be the true value. However, the error of the mean is smaller than the error of any one measurement and the error in the mean gets smaller with more measurements. The standard deviation of the mean is simply given by the standard deviation of a single measurement divided by the square root of the number of measurements minus one.