Chemistry is a science concerned intimately with quantitative measurements. The presentation and treatment of measurements within the context of a report or experiment has many aspects. Two of these will be discussed here: *significant digits* and *the best value*. Indirectly we will need to look briefly at two other important ideas: *accuracy* and *precision*.

The treatment of errors is a science in its own right and this discussion will be brief and somewhat limited. However, the examples to be given are practical for the scope of this course and the student is REQUIRED to employ whatever methods mentioned here may be appropriate to a particular experiment.

I. Significant digits

A common technique for representing uncertainty in a measurement is to record it so that only the final digit (as read from left to right) is "uncertain". For example, a mass measurement from the analytical balance might have an uncertainty as small as 0.2 milligrams (0.0002 g). Masses measured on such a balance should not have any digits in the 0.01 milligram place or smaller. In contrast, typical top-loading balances have an uncertainty of 0.01 grams. Masses recorded from such balances should not have digits in the 0.001 g place or smaller. Representing measurements based upon the uncertainty of the measuring device is an expression of the *precision* of the measurement. A measurement is *precise* if it is <u>repeatable</u> using the same instrument. The number of significant digits in a measurement is a convenient (if approximate) indicator of the probable precision.

The question of "how many significant digits does this measurement have?" is related to the quality of the measuring instrument (and to some smaller extent, the care with which it is used). It is therefore important to know the uncertainty associated with a particular measuring device, whether it is an analytical balance or a graduated cylinder. Generally this information is supplied by the manufacturer. IT SHOULD NOT BE ASSUMED BASED UPON AN ARBITRARY READING OF THE SCALE DIVISIONS OF AN INSTRUMENT.

In order to facilitate further calculations of experimental error, it is suggested that the student form the habit of recording (perhaps at the top of a table entry) the uncertainty of the instrument, generally expressed as ± some value. To that end, the table on the following page includes most of the commonly used lab equipment and the uncertainty associated with it. The significant digits used when recording a measurement are in turn based on these values. For example, a measurement with a standard laboratory thermometer should be written to the nearest tenth of a degree since the uncertainty is $\pm 0.2^{\circ}$C.

Adapted from: <u>Elementary Quantitative Analysis: Theory and Practice</u>, W. J. Blaedel and V. W. Meloche
  <u>Error Analysis</u>, http://science.widener.edu/svb/stats/error.html, Scott Van Bramer
  <u>Error Analysis</u>, http://teacher.nsrl.rochester.edu/phy_labs/AppendixB/AppendixB.html
  <u>Errors, Reliability, and Statistics</u>, http://www.uoregon.edu/~jlong/fa00/237/statistics.html, J. W. Long,
  University of Oregon

Instrument	Uncertainty ($\pm$)
Balance, top loading	0.01 g
Balance, analytical	0.0002 g
-10°C to 110°C thermometer	0.2°C
10 mL graduated cylinder	0.1 mL
25 mL graduated cylinder	0.3 mL
50 mL graduated cylinder	0.4 mL
100 mL graduated cylinder	0.6 mL
100 mL gas measuring tube	0.2 mL
50 mL buret	0.05 mL
1 mL pipet	0.006 mL
10 mL pipet	0.02 mL
25 mL pipet	0.03 mL
100 mL volumetric flask	0.08 mL
250 mL volumetric flask	0.12 mL

II. Significant digits in operations

If the significant digits in a measurement include only one uncertain digit, then it is easy to interpret the often-stated rule for the propagation of significant figures in addition and subtraction. The addition of atomic masses in the example below illustrates this rule:

$$4 \text{ atoms of H} = 4 \times 1.0080 = 4.0320$$
$$1 \text{ atom of C} = 1 \times 12.011 = \underline{12.011}$$
$$\text{formula mass of CH}_4 = 16.043$$

It would be inappropriate to record the mass of CH_4 as 16.0430 since that value would include TWO uncertain digits, the thousandths place being uncertain in the carbon mass and the ten-thousandths places being uncertain in the hydrogen masses. Thus the retention of the least number of decimal places ensures that the number of significant digits is an accurate reflection of the combination of measurements.

The propagation of significant digits in multiplication or division is more complex. The simple rule is to retain in the answer the least number of significant digits among the factors. This approximates the more precise result, the *relative uncertainty*, to be treated in the following section.

For logarithms (base 10 and natural), the number of digits in a measurement determines the number of decimals in the log. The reverse is also true: the number of digits in an antilog is determined by the decimal places in the log.

Students often ask about rounding when doing a series of calculations. In general, each time a result is reported (i.e., some intermediate answer is a value that is requested), it should be rounded to the appropriate number of significant digits following the rules already described. However, continuing calculations *involving* that number should be based on the *unrounded* value, with rounding being done at the final answer.

III. Errors and Accuracy

The *error* in any measurement may be expressed as the difference between the measured value and the true value. If the difference is zero, the measurement is *accurate*. This definition of accuracy may be applied to significant digits. If the significant digits of a measurement agree with the true value (expressed in however many significant digits) then the measurement is considered accurate. It is important to note here that while the <u>number</u> of significant digits is related to the *precision* of a measurement, it has little bearing on the *accuracy*. Thus a poorly used instrument which delivers 6 significant digits in a measurement will not be helpful in obtaining an accurate value.

A good scientist tries to make measurements that are both precise and accurate. In the discussion of errors it is helpful to define alternate terms for precision and accuracy so that results can be compared quantitatively. For a series of N replicate measurements $(X_1, X_2,....X_N)$ of a single quantity, the **mean** (or average) may be defined as

$$\overline{X} = \frac{X_1 + X_2 +...+ X_N}{N}$$

For each of the measurements (and for the mean) there is a finite difference between the value (X_i) and the true (or accurate) value (μ). This difference is called the **error** of the measurement:

$$\text{Error} = \text{observed value - true value}$$
$$= X_i - \mu$$

It is almost intuitively obvious that $\overline{X}$ is the best estimate of μ that can be obtained from a finite set of measurements. **In general the greater the number of measurements, the closer the estimate and the smaller the error.**

The closeness of the measurements $(X_1, X_2,....X_N)$ to one another, without regard to their closeness to the true value (which, after all, may not be known) is known as the **deviation**, which is related to precision. Deviation may be expressed in a number of ways. The deviation (x_i) of any single measurement from the mean would be:

$$x_i = X_i - \overline{X}$$

The **average deviation** $(\overline{x})$ of the set of measurements is the average of the individual deviations, without regard to sign:

$$\overline{x} = \frac{|x_1| + |x_2| +...+ |x_N|}{N}$$

This expression is gradually becoming obsolete because it is not statistically interpretable. Two better expressions of deviation are the *range* and the *standard deviation*.

The **range** (w) is:

$$w = X_{\text{largest}} - X_{\text{smallest}}$$

and the **standard deviation** (s) may be expressed as:

$$s = \sqrt{\frac{x_1^2 + x_2^2 +...+ x_N^2}{N - 1}}$$

As an estimate of precision for a small number of measurements ($N \leq 4$) s is perhaps most common and easiest to use. Unlike the error, all of these quantities can be calculated whether the true value is known or not.

These error and deviation expressions are *absolute*, that is, they are expressed in the same units as the measurements themselves. *Relative* values for these are sometimes more meaningful since they are based on the magnitude of the quantity being measured:

$$\text{Relative error} = \frac{X_i - \mu}{\mu}$$

$$\text{Relative range} = \frac{w}{\overline{X}}$$

$$\text{Relative standard deviation} = \frac{s}{\overline{X}}$$

The sample set of data below can be used to illustrate the calculation of these various expressions.

Number of the measurement, i	Observed Mass, X_i	Deviation from the Mean, x_i	x_i^2
1	25.1 mg	-0.1 mg	0.01
2	25.6	+0.4	0.16
3	25.2	0.0	0.00
4	24.9	-0.3	0.09

$$\overline{X} = 25.2 \text{ mg} \qquad \overline{x} = 0.2 \text{ mg}$$

$$w = 25.6 \text{ - } 24.9 = 0.7 \text{ mg } \textbf{or relative } w = \frac{0.7 \text{ mg}}{25.2 \text{ mg}} = 0.028 \text{ or } 2.8\%$$

$$s = \sqrt{\frac{0.26}{4 \text{ - } 1}} = 0.29 \text{ mg } \textbf{or relative } s = \frac{0.29 \text{ mg}}{25.2 \text{ mg}} = 0.012 \text{ or } 1.2\%$$

$$\text{Error of the mean} = 25.2 \text{ - } 25.6 = \text{-0.4 mg } \textbf{or relative error} = \frac{-0.4 \text{ mg}}{25.6 \text{ mg}} = \text{-0.016 or -1.6\%}$$

Given a series of measurements on a quantity (like the triplicate results of a titration), the mean is generally more correct than any randomly chosen result. But it should be accompanied by an estimate of precision, preferably s or w.

Standard deviation calculations can become very tedious to perform by hand. Fortunately there are relatively simple alternatives available. *Excel* can do a number of statistical manipulations of data. Using the same example data as above, a simple spreadsheet for determining the standard deviation (s) would look like the figure on the following page.

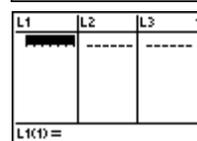|	A	B
1	Observed Mass, mg	
2	25.1	
3	25.6	
4	25.2	
5	24.9	
6	**0.29**	← *s*

The standard deviation is an *Excel* function, STDEV. The formula in cell a6 looks like this:
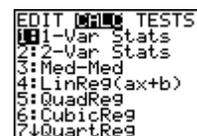
$$=STDEV(A2:A5)$$

The same calculation can also be performed on the TI-83P calculator. The data set must first be placed into L_1. To access the List Editor, press … . The following screen appears:
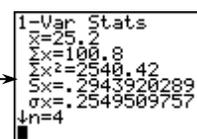
Select EDIT (press 1) for the List Editor:

Type in the four data values under L_1. Exit the List Editor by pressing y z (Quit). Now press … again. Use the arrow keys to select CALC:

Choose 1-Variable Stats (press 1). Press ⌠ :

Among the list of information that appears is the standard deviation (*s*), represented as Sx.

Measurement errors are unavoidable. The many types of errors possible might be divided into two types: *random* and *determinate*. **Random error** varies nonreproducibly from one measurement in a set to another and is not the same magnitude for any two measurements (except by chance). Truly random errors are best addressed by statistical treatments such as those already described. Since random errors vary in both sign and magnitude, they often tend to cancel out if enough samples are done. Repetition of work is the best way, therefore, to minimize the effects of random errors on accuracy (hence the usual procedure of triplicate titration).

Determinate error affects each measurement in the same way and to the same extent. When the difference between $\overline{X}$ and μ is several times larger than $\overline{x}$, *s*, or *w*, the error is clearly determinate. But when the difference is in the same order of magnitude, statistical tests are again needed. Determinate errors include *systematic error*, *constant error* and *proportional error*. **Systematic error** is always in one direction but of irreproducible magnitude. Such errors can often be attributed to experimental conditions which are difficult to control or to the behavior of a particular instrument. Calibration corrections can be helpful in reducing this type of error. **Constant error**, as the name implies, has a reproducible magnitude from one determination to the next. These are common in analyses involving fixed amounts of some reagents in each determination. Correction for this type of error often involves the use of a "blank" determination.

Proportional error involves constant *relative* error and generally results when some incorrect value for the concentration of a solution or similar is assumed in the interpretation of results. This type of error can be accounted for by using controls with known values alongside the unknown determinations.

IV. "The Best Value": the *t-test*

The estimation of the reliability of a derived result is at the heart of error analysis and its correct interpretation allows the reporting of "the best value". Our discussion divides here into two: first the treatment of calculated results derived from multiple trials, and second the treatment of simple calculated results based on measurements from a single instrument or based on a single trial.

In the case of multiple trials, statistical theory provides a method for setting up "confidence limits" for a result (i.e., the probability that the true value, μ, is close to the mean, $\overline{X}$). The confidence limit is a range of values symmetric about the mean. For a small number of trials for which the **true standard deviation** (σ) is unknown, but the **standard deviation** (s) can be calculated, the expression for the best estimate of the true value is:

$$\mu = \overline{X} \pm t\left(\frac{s}{\sqrt{N}}\right)$$

where the *t* values are read from the table below.

Limits of the Confidence Interval around $\overline{X}$, within which μ lies

Confidence level, or % probability that μ
lies *within* the stated interval

N	80% t	90% t	95% t	97.5% t	99% t
2	3.08	6.31	12.71	25.5	63.7
3	1.89	2.92	4.3	6.21	9.92
4	1.64	2.35	3.18	4.18	5.84
5	1.53	2.13	2.78	3.5	4.6
6	1.48	2.02	2.57	3.16	4.03
7	1.44	1.94	2.45	2.97	3.71
8	1.42	1.9	2.36	2.84	3.5
9	1.4	1.86	2.31	2.75	3.36
10	1.38	1.83	2.26	2.69	3.25
	20%	10%	5%	2.5%	1%

Error Probability, or probability that μ
lies *outside* the stated interval

For example, suppose the mass of an object is measured three times, the results being 26.10, 26.28 and 26.22 mg. The standard deviation for this set of measurements is 0.09 mg [you should check this!!!]. Using the values in the table above for a 95% confidence interval, the best estimate of the true value would be:

$$\mu = \overline{X} \pm t\left(\frac{s}{\sqrt{N}}\right) = 26.20 \pm 4.30\left(\frac{0.09}{\sqrt{3}}\right) = \mathbf{26.20 \pm 0.23 \ mg}$$

The meaning of this result is that we would expect 95% of all the means determined from a similar series of repeated measurements to fall within the range of 25.97 to 26.43 mg. The range can be narrowed by either taking more measurements (to lower the standard deviation) or by accepting a lower confidence level.

Like standard deviation calculations, *t*-tests can also be tedious to do by hand. There is no comparable function in *Excel* (although a formula could certainly be written using the STDEV function shown earlier) but the TI-83P calculator can perform this test easily. The data to be tested is entered into L_1 as before. Then press … and use the arrow keys to select TESTS:

Option 8 is the correct one (press 8):
[C-level is the confidence level (usually 95% for this course). The black rectangle in the view above is set on the flashing "STATS". "Data" must be selected (steady black background as shown). The data is already in L_1 so that option is not needed.]
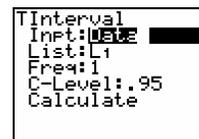
Use the arrow keys to highlight **Calculate** and then press Í :

You can see that the interval matches the values given in the original example. You would need to calculate the ± range yourself.
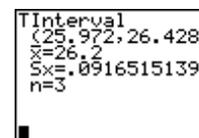
Rejecting data: the Q-test

Sometimes with the limited data sets students collect it is possible that <u>one</u> trial of a multiple set is suspiciously different--or at least seems so--and eliminating it would improve the overall result. Our suspicion is based on possible determinate error. There is an objective statistical test for deciding when a value should be rejected as an "outlier", the *Q-test*.

Using the Q-test involves a series of steps:

1. arrange the data in order (either ascending or descending)
2. calculate the difference (d) between the suspect value and its nearest neighbor
3. calculate the range (r) [the difference between the suspect value and the farthest neighbor]
4. calculate the experimental value for Q_{exp}:

$$Q_{exp} = d/r$$

5. compare Q_{exp} to Q_{table} using the values in the table on the following page

	Confidence level		
	90%	95%	99%
N	Q	Q	Q
3	0.89	0.94	0.99
4	0.68	0.76	0.89
5	0.56	0.64	0.78
6	0.48	0.56	0.78
7	0.43	0.51	0.64

[for $N > 7$, use value for 7]

6. reject the suspect value if $Q_{exp} > Q_{table}$

For results based on a <u>single instrument</u> or a <u>single trial</u> the propagation of error can be based on the *uncertainty* in each measurement as given in the table earlier in this discussion. There are a number of methods for determining the uncertainty of such results, but the simplest may be divided into three approaches---one for additions and subtractions, another for multiplications and divisions and a third for "the rest".

ADDITION AND SUBTRACTION:

For example:
$$12.3013 \pm 0.0002 \text{ g}$$
$$+ \quad 4.2100 \pm 0.0002 \text{ g}$$
$$16.5113 \pm \text{? g}$$

The general approach here may be represented as follows:

$$\text{uncertainty in result} = \sqrt{uncertainty_1^2 + uncertainty_2^2 + ... + uncertainty_N^2}$$

This rule applies equally to addition or subtraction. The uncertainty in the mass represented in the example is thus ± 0.0003 [you should check this for yourself!]

MULTIPLICATION AND DIVISION:

Consider the density calculation below:

$$\text{mass} = 2.31 \pm 0.01 \text{ g} \qquad\qquad \text{volume} = 2.12 \pm 0.05 \text{ cm}^3$$

$$\text{density} = 2.31 \text{ g}/ 2.12 \text{ cm}^3 = 1.09 \pm \text{? g/cm}^3$$

The general approach here may be represented as follows:

$$\text{uncertainty in result} = calculated\ result \sqrt{\left(\frac{uncertainty_1}{measurement_1}\right)^2 + \left(\frac{uncertainty_2}{measurement_2}\right)^2 + ... + \left(\frac{uncertainty_N}{measurement_N}\right)^2}$$

Applying this rule to the problem above we would have:

$$\text{uncertainty in result} = 1.09\sqrt{\left(\frac{0.01}{2.31}\right)^2 + \left(\frac{0.05}{2.12}\right)^2} = 0.03$$

And so the reported density would be 1.09 ± 0.03 g/cm^3.

If standard deviation calculations are tedious these certainly rate some new kind of adjective progressing significantly in the same vein. Fortunately there is (at the time of this writing) a web site with a calculator that will do these sorts of things for you. The URL is **http://web.umr.edu/~gbert/JAVA/uncertainty.HTML** and you can also access the site from the Toolbox area of your own course web site.

Better yet, there is a free program for Windows (XP compatible) which does <u>every</u> test and calculation described in this section. At the time of this writing it can be downloaded from **http://www.shorterfaculty.net/~dturner/downloads/index.htm**

THE REST

The expressions for the less frequently used exponentiation and log functions are given below without further discussion.

EXPONENTIATION as in $x = a^b$ $\qquad$ uncertainty in result $= xb\left(\dfrac{uncertainty_a}{a}\right)$

LOG BASE 10 as in $x = \log a$ $\qquad$ uncertainty in result $= 0.434\left(\dfrac{uncertainty_a}{a}\right)$

ANTILOG as in $x = 10^a$ $\qquad$ uncertainty in result $= 2.303 \cdot x \, (uncertainty_a)$

NATURAL LOG as in $x = \ln a$ $\qquad$ uncertainty in result $= \dfrac{uncertainty_a}{a}$

ANTILN as in $x = e^a$ $\qquad$ uncertainty in result $= x \cdot (uncertainty_a)$

VI. Summary

In brief, all measurements have associated with them a certain amount of error. It is the task of the experimenter to represent this error in a straightforward manner so that his or her data may be interpreted as no more and no less than it actually is. To that end, several methods of various difficulty may be employed. For routine work significant figures are adequate. For formal work, methods involving uncertainties and their propagation through calculations are more appropriate.

In *this course*, the method by which errors are to be tabulated will be described specifically for each experiment. This section thus constitutes a reference for methods of error analysis by particular means.