Abstract
For any quantitative experiment, the basic procedure is to record the data in numbers. In reality no value can be measured with absolute accuracy. For any recorded value, there is always some accompanying uncertainty. In this short script some general guideline is given to treat such real values.

1 Measurement and error
Measurement in reality has always certain uncertainty. Generally the difference of the measured value and the real value is called the absolute error. This means simply:

$$\Delta x = x - \mu$$ (1)

, where $\mu$ is the real value. The error is commonly reported also as relative error $\Delta x/x$ in percentage. All these are quite innocent if just considering the definition. However, in reality we never know the real value without measuring, while measurement always brings in certain error. The error analysis should base on the obtained data, but not simply comparing with theoretical value or reference ata.

Is this correct?
In one of the experiment the student is asked to measure what is the boiling point of water. In this case he obtained 101.4 °C from the thermometer. Than he reports that the error of the measurement is 101.4 - 100 = 1.4 °C, and a relative error of 1.4 %. Does this make sense?

Generally we distinguish two types of error. The random error is simply due to chance. It has no particular pattern and should cancel out if we take the average value of many repeated measurements. The random error can be caused by several reasons. In all measurement setup the measuring instrument has limited precision. This gives the minimum possible error of the measurement, namely the instrumental error. Normally the instrumental error is taken as one or half of the smallest scale. For example, a normal ruler is estimated to have 0.5 mm instrumental error, while a caliper usually have 0.05 mm. In some cases the measurement value will fluctuate and gives different measurement value in independent measurements. In such case the error can be estimated by calculating the standard deviation of multiple measurement values. An example of such procedure is given below.

Another kind of error is the systematic error, which gives a repeatable pattern. Often the cause of the systematic error can be identified and the deviation can be remedied. There is no general method to treat systematic error and each system should be analyzed peculiarly.
Example of random error: repeated measurement

The voltage of a AAA battery is measured 10 times with digital voltmeter. The precision of the voltmeter is 0.01 mV. The measured values are as follow:

1.21941 1.21905 1.21916 1.21909 1.21887 1.21913 1.21892 1.21917 1.21898 1.21919

The unit of the recorded value is in volt. The average of the 10 numbers are taken as the measurement result:

\[ \bar{x} = \frac{1}{N} \sum_{i} x_i = 1.219097 \] (2)

The sample standard deviation of the dataset is defined as:

\[ s = \sqrt{\frac{1}{N-1} \sum_{i} (x_i - \bar{x})^2} = 0.000158 \] (3)

The sample standard deviation tells how much the data points are “spread-out”. This gives the degree of the random error presented in the measurement. However, this is not the error related to the average value that we calculated before. The standard deviation tells more or less how much the value could differ from an expected real value for every measurement data. As can be seen from its formula, the standard deviation will always preserve certain value no matter how much measurement is carried out, but we expect that the more measurement we performed, the better the average should approach reality. The error of the average value is estimated using standard error:

\[ SE_{\bar{x}} = \frac{s}{\sqrt{N}} = 0.000049 \] (4)

This gives the statistic error in the measurement, which is larger than the instrumental error. Because the highest precision is only up to 0.01 mV, one should only preserve 5 digits after decimal point. So the final result of the measurement should be reported as:

1.21910 ± 0.00005 V

One should notice that if we just keep repeating the measurement and the standard deviation maintains a finite value, the standard error will approach 0. This does not mean that one can achieve arbitrary precision because of the finite instrumental error. When the standard error is smaller than the instrumental error, it does not make sense to repeat the measurement again. In this example, one can estimate that after 250 repeating the standard error would be smaller than the instrumental error. Since the standard error decrease inversely to \( \sqrt{N} \), only the first a few repeating measurement would decrease the standard error significantly.

Example of systematic error: two points calibration

A typical systematic error in many measurement system can be identified as a linear shift, expressed as:

\[ \mu = a + b \cdot x \] (5)

while \( \mu \) is the real value and \( x \) is the measured value. The \( a \) and \( b \) are calibration factors. For example, we have a thermometer which measures the mixture of water and ice to be 1.4 °C, and boiling water to be 98.5 °C. With another calibrated thermometer these two temperatures are measured to be 0.2 °C and 100.3 °C. Applying (5) we have \( \mu = -1.243 + 1.031 \cdot x \).

One should not mix up errors with mistakes during the experiment. Usually the mistakes give
strange values and the results are often not reproducible. Although it sounds quite obvious, such error can be very difficult to detect and correct. Visualizing the data will usually help to find out the mistakes. Usually some abrupt changes or abnormality give hint to possible mistakes.

2 Propagation of error

Usually measured values are used in certain formula to deduce other values, and error in the measured values will propagate to the deduced values. Taking measured values to be \( x_i \), the deduced value \( y = f(x_1, x_2...) \), then effect of a small change in \( x_i \) to \( y \) can be calculated using Taylor expansion:

\[
\Delta y = \sum_i \left( \frac{\partial f}{\partial x_i} \right) \Delta x_i + ...
\]  

(6)

Here the higher order terms are neglected. Since the derivative term in (6) can be both positive and negative, but normally the error of \( x_i \) will only propagate positively to \( y \), we should take the absolute value of the derivatives, so the propagation of error is calculated as:

\[
\Delta y \approx \sum_i \left| \frac{\partial f}{\partial x_i} \right| \Delta x_i
\]  

(7)

**Example of error propagation: comparing two large numbers**
The electrical potential of two points in the circuit is measured with respect to the ground. The values are \( 13.475 \pm 0.001 \) V and \( 13.632 \pm 0.001 \) V. Now the voltage drop between these two points is \( 0.157 \) V. The propagated error according to (7) is \( 0.002 \) V. Notice that the directly measured value has 0.007% relative error, but the calculated value has 1.3% relative error, which is about 2 order of magnitude higher. This example shows that when comparing a small difference between two large values, because of the finite precision the result might have very large error.

The error calculated according to (7) gives more or less the maximum deviation. In reality the possibility that all the \( \Delta x_i \) take the maximum deviation is quite low. If we think \( \Delta x_i \) as random variable, and (7) as a combination of random variables, the total variance of \( y \) can be expressed as:

\[
\sigma_y^2 = \sum_i (\frac{\partial f}{\partial x_i})^2 \sigma_{x_i}^2 + \sum_{i,j} (\frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j}) \text{COV}_{x_i,x_j}
\]  

(8)

, where \( \text{COV}_{x_i,x_j} \) is the covariance of \( x_i \) and \( x_j \). If all the \( \Delta x_i \) are independent to each other, the covariance terms are 0.

3 Correlation and least square fitting

Two measurement values are correlated if the change of one variable causes a defined and repeatable change of another one. There are many method to test the correlation between two variables. The most simple way is to plot the data points, as is shown in figure 1.

**Correlation and causation**
Correlation does not necessarily means there is a relation of causality between two variables. Two variables showing strong correlation can bear no physical significance at all. An example is shown in figure 2. To identify that change of something causes change of another, one needs careful analysis.

If two variables are correlated, the relation can be expressed in certain formula with some parameters. The fitting or regression is the procedure to find parameters for certain formula that suits the
data points. The least mean square method is a common method to perform the fitting process. Assume that variable $x$ and $y$ are correlated by formula $y = f(x; a_1, a_2...)$, where $a_1, a_2...$ are the fitting parameters. The measurement data are $(x_i, y_i)$. The total deviation of the data to the fitting curve is:

$$I = \sum_i \left[ y_i - f(x_i; a_1, a_2...) \right]^2 \quad (9)$$

The best fitting parameters are those having the smallest total deviation $I$. The task of fitting is then transformed to minimization of $I$ according to $a_1, a_2...$. This means also to solve the following equations for $a_1, a_2...$:

$$\frac{\partial I}{\partial a_j} = -\sum_i \frac{\partial f(x_i; a_1, a_2...)}{\partial a_j} \cdot \sum_i \left[ y_i - f(x_i; a_1, a_2...) \right] = 0 \quad (10)$$

Generally, equation (10) cannot be solved algebraically, but numerical method is applied to minimize (9) directly. The detailed procedure is very complicated and will not be covered here. The basic idea is that we start from an initial guess of the parameters, and then changing the parameters in small steps and find out the change in which reduce the value of $I$. These small steps are applied recursively to the parameters until finally $I$ reaches the minimum. The whole method is basically a guided try-and-error process. The function $I$ might have lots of local minimum values in respect to $a_j$,
normally the fitting procedure can only find the local minimum which is closest to the initial values of \( a_j \). Such procedure is implemented by many programs and can be readily used.\(^1\)

**Example of fitting: linear fit**

If \( y = f(x; a_1, a_2...) \) is a linear equation, (10) can be solved algebraically. Such process is also called linear regression. Take \( y = a + bx \), the best fitting parameters are:

\[
\begin{align*}
  b &= \frac{\sum_i^N x_i y_i - N \bar{x} \bar{y}}{\sum_i^N x_i^2 - N \bar{x}^2} \\
  a &= \bar{y} - b \bar{x}
\end{align*}
\]

(11) (12)

If the measurement error in \( x \) and \( y \) are purely statistic, the variance of \( x \) and \( y \) can be calculated as:

\[
\begin{align*}
  \sigma_x^2 &= \frac{\sum_i^N x_i^2 - N \bar{x}^2}{N} \\
  \sigma_y^2 &= \frac{\sum_i^N y_i^2 - N \bar{y}^2}{N} \\
  \text{COV}_{x,y} &= \frac{\sum_i^N x_i y_i - N \bar{x} \bar{y}}{N}
\end{align*}
\]

(13) (14) (15)

and the standard error of \( a \) and \( b \) is:

\[
\begin{align*}
  SE_a &= s \sqrt{\frac{1}{N} + \frac{1}{\sum_i^N x_i^2 - N \bar{x}^2}} \\
  SE_b &= s \sqrt{\frac{1}{\sum_i^N x_i^2 - N \bar{x}^2}}
\end{align*}
\]

(16) (17)

, where \( s \) is defined as:

\[
\begin{align*}
  s &= \sqrt{\left[ \sum_i^N y_i^2 - N \bar{y}^2 - \frac{(\sum_i^N x_i y_i - N \bar{x} \bar{y})^2}{\sum_i^N x_i^2 - N \bar{x}^2} \right] / (N - 2)}
\end{align*}
\]

(18)

\(^1\)Several programs that perform least square fitting:

- Origin (Complete data analysis program for Windows) [http://www.originlab.com/](http://www.originlab.com/)
- gnuplot (open source program for Windows, Linux) [http://www.gnuplot.info/](http://www.gnuplot.info/)