MEASUREMENTS AND ERRORS

STUDY GUIDE

For PHYSICS LABS

INTRODUCTION:

This guide is summary of the basic concepts, definitions and formulas involved in measurement, statistical treatment of data, and combination of errors.

TOPIC: BASIC MEASUMENT CONCEPTS

Direct Measurements:

A direct measurement is physical comparison of unknown quantity to a known amount of the same kind of quantity. For example, length of a material object compared to a linear measuring instrument (ruler, meter stick, calipers, etc.). The result of the measurement process is a number represents the total number of the standard units of measure needed to equal the amount in the unknown.

Indirect Measurements:

Measurement may also be indirect. An indirect measurement is a calculated result. Two or more direct measurements of different physical quantities are used to compute the value of another entirely different kind of physical quantity. For example, you can directly observe and measure both "distance" and "time" of a moving object. From these two results, one can then calculate the object's velocity. It's also possible to say one has "measured" the mass of an electron or the mass of the earth, but these are indirect measurements and are operationally quite different measurements from the direct mass measurements made of an object placed on a lab balance. Calculations are involved and it requires that a valid mathematical formula exist.

Error:

The error is the arithmetic difference between the measured and "true" or "accepted" result. The error has the same units as the measurement results, e.g. meters, seconds, grams, volts etc.

Relative Error and Percentage Error:

Rel Error = error / measured quantity

STATISTICAL CONCEPTS—SINGLE VARIABLE

Statistics is an enormous subject and has application not only in physical science but in marketing, politics, social, and education to name just a few areas. What is summarized below is enough to get you started in beginning physics lab. You should consult other texts for more details.

Average and Arithmetic mean: The mean of a series of measurements is obtained by adding all values, x, and dividing by the number of observations, N.

$$\bar{x} = \frac{\sum_{1}^{n} x_i}{N} \qquad \text{Eq. 1}$$

Deviation from the mean: The arithmetic difference of a single observation from the mean value.

$$d_i = x_i - \bar{x} \qquad \text{Eq. 2}$$

Standard Deviation (of a single observation): This is common measure of the spread of a set of data. It is also called the root-mean-square deviation and is calculated as

$$\sigma_N = \sqrt{\frac{\sum (x_i - \bar{x})^2}{N}} = \sqrt{\frac{\sum d_i^2}{N}} \qquad \text{Eq. 3a}$$

Note that when N is used as above, the result is called the "observed standard deviation". For small sets of data, very often the above formula is modified using N-1 in the denominator and this is called…

Estimated Standard Deviation:

$$\sigma_{N-1} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{(N-1)}} = \sqrt{\frac{\sum d_i^2}{(N-1)}} \qquad \text{Eq. 3b}$$

Standard Error of the Mean: This is an estimate of the deviation of mean value itself. The mean of course is the result of a number of observations, N. The standard error is computed from

$$\sigma_M = \frac{\sigma_N}{\sqrt{N-1}} \qquad \text{Eq. 4}$$

Equation (4) agrees with the common sense notion that "mean" of several observations is more reliable than a single observation. Equation 3a expresses the standard deviation of a single observation. It turns out that there will be a 68% probability that any single observation will not differ from the "mean" value by more than one standard deviation. Also, 95% of the single observations will lie within two standard deviations and approximately 99.7% lie with three standard deviations. The point is, there is a probability that a single observation will differ from the mean value. Whereas, equation 4 gives us a way to express the standard error in the mean value itself.

LINEAR RELATIONSHIPS and REGRESSION ANALYSIS

The above definitions and formulas apply to the statistics of a single variable. Next, We want to summarize how one goes about determining the correlation of pairs of data, say variables x and y. The goal is to find out if any variation in the independent variable x is followed by proportional variation in y. This is called a regression analysis. We'll summarize how to do a linear regression analysis, that is find the best straight line to fit a range of variable, x and y.

Graphical Method

It's generally a good idea to plot your data first to see what is going on, to determine if you have data sets spaced over an adequate range, does it look approximately linear, etc. If greater accuracy is desired, then an analytic fit to your data may be justified.

First, select scales for x and y to cover the range of your data that will give you approximately a 45 degree slope, then plot your pairs of points. If you have estimates in the errors in y, also plot these as error bars; vertical lines extending above and below the central y value.

Then use a straight edge to lay out and draw the best straight line through all your data. The "best line" will have an equal number of points above and below it. [Analytically, the best line minimizes the sum of the squared deviations.]

The intercept is the point at which the line crosses the y-axis. The slope is found by taking pairs of value <u>from the line</u> and calculate using:

$$b = \frac{y_2 - y_1}{x_2 - x_1}$$

The errors in the slope and intercept can also be estimated from this graph, at least within the resolution of the scale you selected.

Analytical Method

Here, you want to use a set of data points (x_1, y_1; x_2, y_2; …) and calculate the "best fit" of a straight line to the equation

$$y = bx + a \qquad \text{where b is the slope and a the intercept.}$$

The 'best fit" minimizes the sum of the squares of the deviations of your "y" data points from the straight line. Incidentally, this method and the derived formulas assume that the variation in the individual x values is negligible.

The slope, b. of the best straight line is

$$b = \frac{\sum_i (x_i - \overline{x})(y_i - \overline{y})}{N\sigma_x^2} \qquad \text{Eq. 5}$$

The intercept on the y-axis is

$$a = \overline{y} - b\overline{x} \qquad \text{Eq. 6}$$

The errors in the slope and the intercept are given by

$$\sigma_b = \frac{1}{\sigma_x}\sqrt{\frac{\sum d^2}{N(N-2)}} \qquad \sigma_a = \sigma_b\sqrt{\sigma_x^2 + \overline{x}^2} \qquad \text{Eq. 7}$$

where d in equation 7 means $d_i = y_i - (bx + a)$; the difference between the point, y_i, and the calculated value (bx + a).