

KINDS OF MEASUREMENTS AND THEIR RELIABILITY

One aspect of physical sciences and engineering which differs from most other disciplines is the prevalence of quantitative measurements. Observations in the physical sciences often generate measurements, i.e., numbers. In most cases, a measurement can be of either a single property, or it can involve two properties where the property being measured depends in some way on the other property.

Single property measurements are normally repeated several times. While in principle one would expect identical values among a set of measurements, in practice there are many factors which cause the repeated measurements to be different from each other. A common example from chemistry is the titration to an endpoint of identical samples. In principle, the same volume should be required for reaching the endpoint in each trial, but in practice, there will be some variance in the volumes.

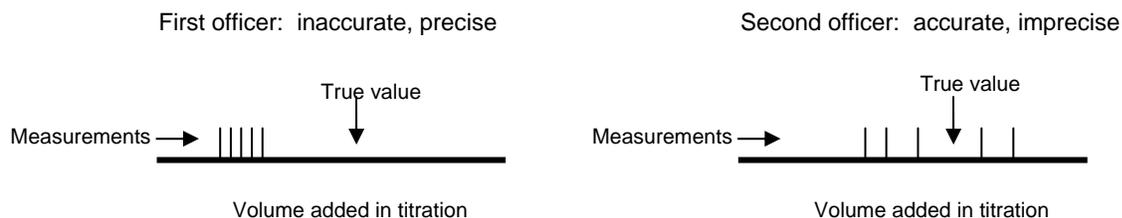
Measurements involving the relationship between two properties are normally carried out by measuring the value of one of the variables, then measuring the other, and then repeating this several times for different values of the first variable, resulting in a series of pairs of measurements. An example is the measurement of the mass of a fluid as the volume of the fluid is varied through several values. In principle, a consistent pattern of dependence of the one variable on the other is expected, but in practice, there will be some variance from absolute consistency. The goal of such a measurement process is to determine the exact way in which the second variable (e.g., mass) depends on the first variable (e.g., volume). In many cases, as in the example of mass vs. volume, the graphed points produced by plotting the pairs of measurements follow a linear trend. In the case of mass vs. volume, the slope of this line has physical significance, e.g., it corresponds to the density of the fluid.

Accuracy and *precision* are two different terms that are used to refer to the numbers that result from measurements, and both are applicable to single property and two-property measurements. *Accuracy* refers to the agreement between the measured quantity and the “true” value. An accurate value is one that is very close to the true or actual value. *Precision* refers to the degree of consistency among a set or series of measurements. It reflects the reproducibility of the measurements.

Single Property Measurements

Consider two junior officers learning about the chemical tests performed on boiler water, several of which involve specific titrations to determine the amount of certain chemical species in the water. They are each provided several identical samples of boiler water, and perform the same chemical titration on those samples. As each titration is completed, they record the volume of solution added. This volume represents a single property, and in principle, one would expect identical values for each trial performed by the two officers.

The results for the two officers’ trials are symbolically represented below. Clearly, each trial did not result in an identical measurement. In fact, there are some glaring differences which warrant further explanation and clarification.



The first officer's measurements were very close to each other, but none of them was close to the true value. The second officer's measurements were not very close to each other, but they tended to surround the true value. What do these results mean in terms of *accuracy* and *precision*? The first officer's measurements, though highly precise, are clearly very inaccurate. The second officer's measurements, though not as precise, yield an average which is much more accurate (i.e., closer to the true value).

There is some inherent *systematic error* (i.e., always too high or always too low) in the first officer's measurements, even though the way he or she performs the measurements is highly reproducible from trial to trial. Systematic error can arise for a variety of reasons (e.g., burets with misplaced markings, or consistently improper reading of the meniscus). One must carefully account for and explain systematic error. Generally precautions are taken to avoid such error, if within the control of the investigator.

Both officers' readings, as is the case for all measurements, include *random error*, where measurements fall slightly above or below the average. The first officer's measurements have smaller random errors than the second officer's measurements. This relates to the precision.

Random Error Analysis

A single measurement of some property (e.g., mass or volume) may be randomly either slightly too high or too low. If the property is measured several times, and then averaged, these random deviations will tend to cancel out. Thus, more measurements of the property followed by averaging should give a better (i.e., more accurate) result for the measurement, so long as there are no *systematic errors* inherent in the measurement process. The precision of the measurement is related to the size of the random deviations of the individual measurements from the average. Two common methods for determining precision are (1) the average deviation from the mean, and (2) the standard deviation from the mean.

Example of Error Analysis

Actual data from an officers' titrations will be used below to illustrate the terms and methods described above. The officer performs four titrations on identical boiler water samples, resulting in titration volumes of 6.58 mL, 6.66 mL, 6.61 mL, and 6.55 mL.

A. Mean or average

Often, the first step in analyzing data is to determine the mean or arithmetic average. In general, the actual or true value lies within the range of the measured values if there is no significant systematic error. The mean (\bar{x}) is calculated by dividing the sum (Σ) of the values obtained (x_i) by the total number of values (n):

$$\bar{x} = \frac{\sum x_i}{n}$$

$$\text{Example: } (6.59 \text{ mL} + 6.66 \text{ mL} + 6.61 \text{ mL} + 6.55 \text{ mL}) / 4 = 6.60 \text{ mL}$$

B. Average deviation from the mean

The average deviation from the mean is the simplest method of describing precision. It indicates how close, on an average, the individual values are to the average or mean. A set of

measurements with a large average deviation from the mean is not very precise. The average deviation from the mean is calculated by taking, for each data point, the absolute value of the difference between a data point and the mean, and then averaging these:

$$\text{Average deviation from the mean} = \frac{\sum |x_i - \bar{x}|}{n}$$

$$\text{Example: } (0.01 \text{ mL} + 0.06 \text{ mL} + 0.01 \text{ mL} + 0.05 \text{ mL}) / 4 = 0.03 \text{ mL}$$

Therefore, using the mean and the average deviation from the mean, the reported titrated volume would be $6.60 \pm 0.03 \text{ mL}$.

C. Percent average deviation from the mean

The percent average deviation from the mean compares the average deviation from the mean to the mean itself. For example, a one meter average deviation in measuring the distance to the moon gives a very small percent average deviation from the mean because the distance measured is so much greater than the average deviation. When shooting basketballs, however, a one meter average deviation from the basket gives a very large percent average deviation from the mean because the distance measured is comparable in magnitude to the average deviation. The % average deviation from the mean is calculated by taking the percent of the average deviation from the mean and dividing it by the mean:

$$\text{percent average deviation from the mean} = \left[\frac{(\sum |x_i - \bar{x}|) / n}{\bar{x}} \right] \times 100\%$$

$$\text{Example: } [0.03 / 6.60] \times 100\% = 0.45\%$$

D. Standard deviation from the mean

The standard deviation from the mean is commonly used when there are many individual measurements being analyzed. It is similar to the simple average deviation from the mean (see B above), except that it involves the *squares* of the individual deviations rather than the absolute deviations. It is defined as follows:

$$\text{standard deviation from the mean} = \left(\frac{\sum (x_i - \bar{x})^2}{n - 1} \right)^{1/2}$$

where n is the number of measurements. If the “n-1” term in the above expression is replaced by “n,” the expression would be simply interpreted as the square root of the average of the squared deviations from the mean. For theoretical reasons, however, the term “n-1” is employed instead of “n.” This method of determining the precision of a measurement process is extremely useful in many applications of statistical analysis.

$$\text{Example: } [(0.01^2 + 0.06^2 + 0.01^2 + 0.05^2) / 3]^{1/2} = 0.05 \text{ mL}$$

E. Absolute error

Absolute error is a measure of the accuracy of a measurement. It is a comparison of the measured value and the true or accepted value. It is defined as the difference between the observed value and the accepted value, x_i . The concept can be applied to individual measurements or to averages of measurements:

$$\text{Absolute error (for an individual measurement)} = x_i - x_t$$

Example: If the “correct” value for the titration were known to be 6.62 mL, the first titration volume would have an absolute error of $6.58 \text{ mL} - 6.62 = -0.04 \text{ mL}$. The “sign” is included, providing explicit information about the direction of the error (i.e., too high or too low).

F. Relative error

Relative error (or percent error or percent deviation) is often a more useful quantity than absolute error. The relative error is the percentage of the absolute error (for an individual measurement or for an average of measurements) divided by the accepted value:

$$\text{percent error} = [(x_i - x_t)/x_t] \times 100\%$$

Example: The first titration volume would have a percent error of $(6.58 - 6.62)/6.62 \times 100\% = -0.60\%$

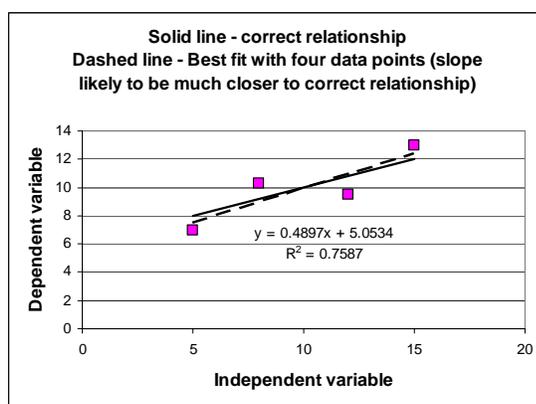
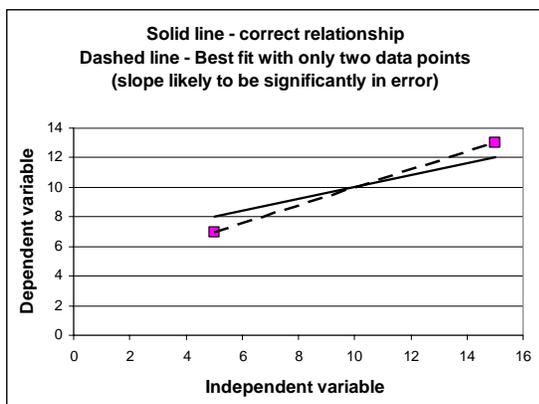
Independent and Dependent Property Measurements (Two Variables)

The same concepts also apply to measurements of some property (e.g., mass) which depends on another variable (e.g., volume). Typically, the *dependent* property is measured for each of a series of values of the *independent* variable (e.g., masses measured for each of a series of volumes). For this situation, the data are usually plotted on a graph, with the *dependent* property as the y-coordinate, and the *independent* variable as the x-coordinate. For cases where the plotted data points follow a linear trend, the slope of the line is often of importance (e.g., in the case of the mass and volume, the slope would be the density = mass/volume).

Systematic error can occur in this type of analysis, just as in the single property measurements discussed above. Either or both of the measured variables can be subject to systematic error. As before, one must carefully account for and explain systematic error. Generally precautions are taken to avoid such error, if within the control of the investigator. In the discussion below, we will assume that there is no systematic error in the measurement process.

Error Analysis for Linear Dependence Case

If only two data points are obtained, each point being randomly slightly too high or too low, the slope of the line through the points will likely be in error. If several data points are obtained, however, the random highs and lows of each measurement tend to cancel each other out more effectively, and the best-fit line through the points will be closer to the correct value (i.e., more accurate), as illustrated in the following graphs:



The precision of this set of measurements is related to how close each individual data point is to the best-fit line through the points. Best-fit straight lines are usually characterized by their slope and intercept values, as indicated on the second graph above. The precision of best-fit lines is usually characterized by the quantity R^2 , also known as the correlation coefficient. The closer this coefficient is to the number 1, the more precise or closer the individual data points come to falling along the straight line. High precision measurements will typically have R^2 values very close to 1, e.g., 0.99 or higher.

Tools for Analyzing Measurements

Spreadsheets have special functions which allow for the determination of simple averages, average deviations, and standard deviations for repeated single property measurements, and slopes, intercepts, and quality-of-fit for measurements involving two properties. Some calculators are also pre-programmed with functions that yield these values. Students should become familiar with these calculational tools and use them routinely in analyzing laboratory data.

Bad Data Points

Sometimes, a data point among a set of repeated measurements is obviously in error for some reason. Equivalently, a point among a set of plotted points may be clearly in error compared to most or all of the other points. Though there are some sophisticated statistical methods for correctly determining whether to reject such points from your data sets, at this level of treatment, students should feel empowered to intuitively decide whether to remove such points. If omission of a data point reduces the total number of usable points to too small a number (e.g., two), it is best to perform the measurement again, if possible.

