~° Error °~

No matter how carefully a quantitative determination is performed, the result always differ to some extent from the true content i,e it contains certain error.

Therefore error may be defined as "the numerical difference between a measured value and the absolute or true value of an analytical ~~measurement~~ determination." The absolute or true value of a quantity is, however, never known. All that we can use is only an "accepted value". The value for any quantity is "accepted" when the only uncertainty in this value is less than the uncertainty in some other quantity with which the other quantity is to be compared. Error often denotes the estimated uncertainty in a measurement or experiment.

Classification: Analytical errors can be classified as

(a) Systematic or determinate errors
(b) Random or indeterminate errors.

Systematic error:

Systematic errors are errors which are constant in magnitude or vary in accordance with a definite law. They have an assignable cause and are of same magnitude for replicate measurements in the same way. A given determinate error is generally unidirectional with respect to the true value and thus makes measured value either lower (-ve bias) or higher (+ve bias) than the true value. The following types of systematic errors may be noted:

(a) Methodic error —

These errors depends on specific characteristics of the analytical method used. e.g., the reaction on which the determination depends may not be quite quantitative; the precipitate may be partially soluble, or various impurities may be precipitated with it, the precipitate may partially decompose or volatilize during ignition, the ignited precipitated may be hygroscopic; the main reaction may be accompanied by side reactions which distort the results of the determinations.

These errors generally arise from non-ideal chemical or physical behaviour of the reagents and reactions on which an analysis is based. Such sources of

nonideality include the slowness of some reactions, the incompleteness of others, the instability of some species, the non specificity of most reagents, and the possible occurance of side reactions that interfere with the measurement process.

The errors inherent in a method are often difficult to detect and thus most serious cause of incorrect results in quantitative determinations.

(b) Instrumental error —

These errors arise from the imperfections in measuring device. e.g. pipets, burettes, and volumetric flasks may hold or deliver volumes slightly different from those ~~idi~~ indicated by their graduations. These differences arise from using glassware at a temperature that differ significantly from the calibration temperature, from distortions in container walls due to heating while drying, from errors in the original calibration, or from contaminants on the inner surfaces of the ~~containers~~ containers. Calibration eliminates most systematic errors of this type.

Electronic instruments are subject to instrumental systematic error. These can have many sources e.g. (i) Errors may emerge as the voltage of a battery-operated power supply decreases use.

(ii) Errors can also occur if instruments are not calibrated frequently or calibrated incorrectly.

(iii) The experimenter may also use an instrument under conditions in which errors are large.

(iv) Temperature changes cause variation in many electronic components, which can leads to drifts and errors. Such instruments are susceptible to noise induced from the alternating current power lines, and this noise may influence precision and accuracy.

In many cases, errors of these types are detectable and correctable

(c) Operative errors — Operative errors are due to the faulty or careless execution of analytical operations. Operations in which these errors may occur include transfer of solutions, effervescence and 'bumping' during sample dissolution, incomplete drying of samples, underwashing/overwashing of

precipitate and so on.

(d) Personal errors – These errors depend on the personal characteristics of the analyst himself. e.g. his inability to detect exactly the end-point in titration; etc. Personal errors also include so-called psychological errors, due to a certain bias often met with in students, e.g. in duplicate weighings or titrations, out of two adjacent scale divisions on the balance or burette, the student often tends to choose the division which is closer to the previous determination or even to those found by his fellow students rather than the one closer to the actual weight or volume.

Minimisation of systematic error:

Systematic or determinate errors can often by materially reduced by one of the following methods:

(a) Calibration of apparatus and application of corrections:- All instruments (weights, flasks, burettes, pipettes etc) should be calibrated and appropriate corrections applied to the original measurements. In some cases, where an error can not be eliminated, it is possible to apply a correction for the effect that it produces; thus an impurity in a weighed precipitate may be determined and its weight deducted.

(b) Running a blank determination – This consists in carrying out a seperate determination, the sample being omitted, under exactly the same experimental conditions as are employed in the actual analysis of the sample. The object is to find out the effect of impurities introduced through the reagents and vessels, or to determine the excess of standard solution necessary to establish the end point under the conditions met within the titration of the unknown sample. A large blank correction is undesirable, because the exact value then becomes uncertain and the precision of the analysis is reduced.

(c) Running a control determination — This consists in carrying out a determination under as nearly as possible identical experimental conditions upon a quantity of a standard substance which contains the same weight of the constituent as it contained in the unknown sample. The weight of the constituent in the unknown can then be calculated from the relation:

$$\frac{\text{Result found for standard}}{\text{Result found for unknown}} = \frac{\text{weight of constituent in std}}{x}$$

where x is the weight of the constituent in the unknown.

(d) Use of independent methods of analysis — In some cases the accuracy of a result may be established by carrying out the analysis in an entirely different manner. Thus Iron may first be determined gravimetrically by precipitation as Iron(III) hydroxide after removing the interfering elements, followed by ignition of the precipitate to Iron(III) oxide. It may then be determined ~~titri~~ titrimetrically by reduction to the iron (II) state, and titration with a standard solution of an oxidising agent, such as $K_2Cr_2O_7$. If the result obtained by the two radically different methods are concordant, it is highly probable that the values are correct within small limits of error.

(e) Running of parallel determinations — These serve as a check on the result of a single determination and indicate only the precision of the analysis. The values obtained for constituents which are present in not too small an amount should not vary among themselves by more than 3 ppt. If larger variations are shown, the determinations must be repeated until satisfactory concordance is obtained. Duplicate / triplicate determinations should suffice. It must be emphasised that good agreement between duplicate or triplicate determinations does not justify the conclusion that the result is correct,

a constant error may be present. The agreement nearly shows that the accidental errors, or variations of the determinate errors, are the same, or nearly the same, in the parallel determinations.

(f) Standard addition— A known amount of the constituent being determined is added to the sample, which is then analysed for the total amount of constituent present. The difference between the analytical results for samples with or without the added constituent gives the recovery of the amount of added constituent. If the recovery is satisfactory, our confidance in the accuracy of the procedure is enhanced. The method is usually applied to physico-chemical procedures such as polarography and spectrophotometry.

(g) Internal standards— It involves adding a so fixed amount of a reference material (internal standard) to a series of known concentrations of the material to be measured. The ratio of the physical value (e.g. absorbance or peak size in spectrophotometry or chromatography respectively) of the internal standard and the series of known concentrations is plotted against the concentration values. This should be a straight line. Any unknown concentration can then be determined by adding the same quantity of ~~unknown~~ internal standard and finding where the ratio obtained falls on the concentration scale.

(h) Amplification methods— In determinations in which a very small amount of material is to be measured, this may be beyond the limits of the apparatus available. In these circumstances if the small amount of material can be reacted in such a way that every molecule produces two or more molecules of some other measurable material, the amplification of the quantity may then be within the scope of the apparatus or method available.

(i) Isotope dilution — A known amount of the element being determined, containing a radioactive isotope, is mixed with the sample and the element is isolated in a pure form (usually as a compound), which is weighed or otherwise determined. The radioactivity of the isolated element is measured and compared with that of the added element; from this the weight of the element in the sample can then be calculated.

Constant errors:

Determinate or systematic error may be classified as constant or proportional.

A constant error is independent of the magnitude of the measured quantity and becomes less significant as the magnitude increases. It also independent of the concentration of the substance being analysed.

e.g. (a) In volumetric analysis, the excess of the titrant (say, 1 drop) that has to be added to bring about a change of colour at the end point remains the same whether we titrate 5 ml or 25 ml of the soln

(b) If a constant end point error of 0.10 ml is made in a series of titrations, this represents a relative error of 1% for a sample requiring 10 ml of titrant, but only 0.2% if 50 ml of titrant is used.

Proportional errors:

The absolute value of this type of error varies with sample size in such a way that the relative error remains constant. A substance that interferes in an analytical method may lead to such an error if present in the sample.

e.g. In the iodometric determination of an oxidant like chlorate, another oxidising agent such as bromate would cause high results if its presence were unsuspected and not corrected for. Taking larger samples would increase the absolute error, but the relative error would remain constant provided the sample was homogeneous.

Accuracy & Precision

Accuracy:

Accuracy indicates the closeness of the measurement to the true or accepted value. It is often expressed in terms of error. For a set of measurements, accuracy indicates the closeness of the mean value of the set of measurements to that of the true or accepted value.

Thus, accuracy = measured value − true value

= Mean value − true value

Smaller the difference between the mean value and the true value, larger will be the accuracy.

Absolute error:

The absolute error is the difference between the experimental value and the true value.

Thus the absolute error, E, in the measurement of a quantity x is given by the equation

$$E = x_i - x_t$$

where x_i is the true or accepted value of the quantity.

The sign of the absolute error tells whether the value in question is high or low. If the measurement result is low, the sign is −ve; if the result is high, the sign is +ve.

Example: If an analyst finds a value of 20.44% iron in a sample which actually contains 20.34%, then calculate the absolute error in the measurement?

Absolute error: $20.44 - 20.34 = 0.10\%$.

Relative error:

The error is most frequently expressed relative to the size of the measured quantity. In that case, the relative error, E_r, is a more useful quantity than the absolute error. The relative error is expressed either in percent or in parts per thousand (ppt). Thus,

$$E_r = \frac{x_i - x_t}{x_t} \times 100\%$$

$$= \frac{x_i - x_t}{x_t} \times 1000 \text{ ppt}.$$

where x_t = true / accepted value for measurement

x_i = Measured value.

Example: In an experiment, the concentration of Zinc in a given sample was found to be 20.17 ppm. Taking the accepted value as 20.00 ppm, calculate the relative error as percent and parts per thousand.

Relative error: $E_r = \frac{20.17 - 20.00}{20.00} \times 100\% = .85\%$

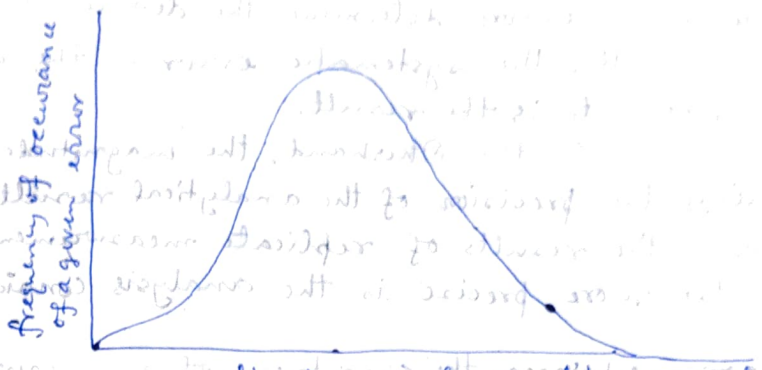
$$= \frac{20.17 - 20.00}{20.00} \times 1000 = 8.5 \text{ ppt}.$$

Precision:

The degree of agreement between two or more replicate measurements made on a sample in an identical manner is known as the precision of the measurement.

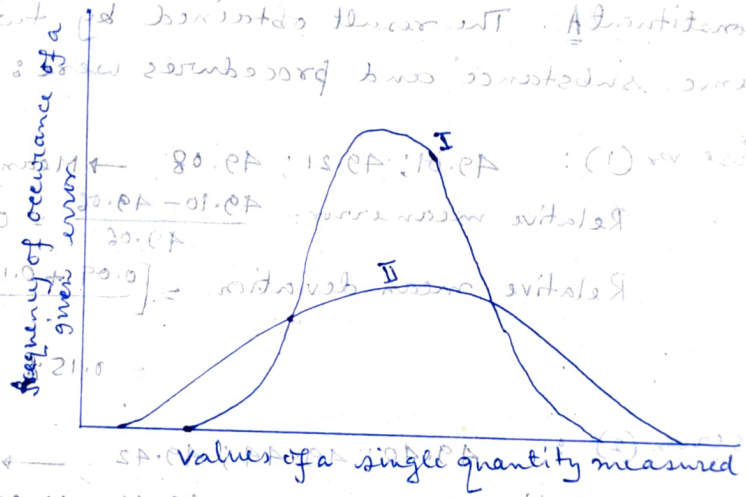
Precision reflects the closeness among replicate measurements; it reflects nothing about their relation to the true value. Precise value may well be inaccurate, since an error causing deviation from the true value may affects all the measurements equally and hence not impair their precision. A determinate error which leads to inaccuracy may or may not affect precision, depending upon how nearly it remains constant through out a series of measurements. The precision is commonly stated in terms of standard deviation, average deviation or range.

If we make a large no. of observations of a single quantity and then plot the no. of times a given value occurs against the value of the quantity itself, we obtain a curve known as ~~Normal~~ error distribution curve as shown



The error distribution curve

These curves have two useful features, viz, the height of the peak of the distribution and the spread of the distribution curve (also called dispersion)



Characteristics of less precise and more precise result

The precision of two set of measurements of the same quantity can thus easily be compared with the help of their error distribution curves. Thus in the above figure, Curve I with less spread and higher peaks is characteristic of more precise measurements as compared to the curve II, which represents less precise measurements.

Difference between accuracy and precision:

(a) Accuracy is the closeness between the measured value and true or accepted value while precision is the degree of agreement between two or more replicate measurements made on a sample in an identical manner.

(b) Systematic errors determine the degree of accuracy of the result; smaller the systematic error in the determination, the more accurate is the result.

On the Otherhand, the magnitude of random errors determines the precision of the analytical result. It follows that the closer the results of replicate measurements are to each Other, the more precise is the analysis considered to be

*→(c) Accuracy expresses the correctness of a measurement, while precision represents the reproducibility of a measurement.

(d) Precision always accompanies accuracy, but a high degree of precision does not imply accuracy.

Example:

A substance was known to contain $49.06 \pm 0.02\%$ of a constituent A. The result obtained by two observers using the same substance and procedures were:

Observer (1): 49.01; 49.21; 49.08 → Mean = 49.10%

Relative mean error: $\frac{49.10 - 49.06}{49.06} = 0.08\%$

Relative mean deviation $= \left[\frac{0.09 + 0.11 + 0.02}{3}\right] \times \frac{100}{49.10}$

$= 0.15\%$

Observer (2): 49.40; 49.44; 49.42 → Mean = 49.42%

Relative mean error: $\frac{49.42 - 49.06}{49.06} = 0.73\%$

Relative mean deviation $= \left[\frac{0.02 + 0.02 + 0.00}{3}\right] \times \frac{100}{49.42}$

$= 0.03\%$

Therefore, the analyses of observer (1) were therefore accurate and precise; those of observer (2) were unusually precise, but less accurate than those of observer (1).

Random / Indeterminate error :

Random or indeterminate, errors are caused by many uncontrollable variables that are an inevitable part of every analysis. Most contributors to random error can not be positively identified, thus they can never be eliminated and are often major source of uncertainty in determination. However if the source of uncertainty is identified, it is usually impossible to measure them because most are so small that they can not be detected individually. The accumulated effect of the individual uncertainties, causes replicate measurements to fluctuate randomly around the mean of the data set, so that both high and low results are equally probable.

Random errors follow a normal distribution and mathematical laws of probability can be applied to arrive at some conclusion regarding the most probable result of a series of measurements. When the normal distribution of indeterminate errors are graphically represented, it known as error distribution curve or normal error curve or Gaussian curve.

A typical error curve is as shown below and it satisfies the following equation

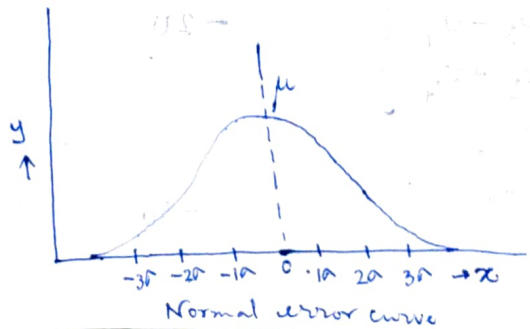
$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where y = Relative frequency of errors

σ = standard deviation of infinite population (known as population standard deviation)

x = One of replicate measurements (numerical value of a replicate measurements)

μ = population mean



Normal error curve

Thus a Gaussian or normal error curve shows the symmetrical distribution of data around the mean of an infinite set of data.

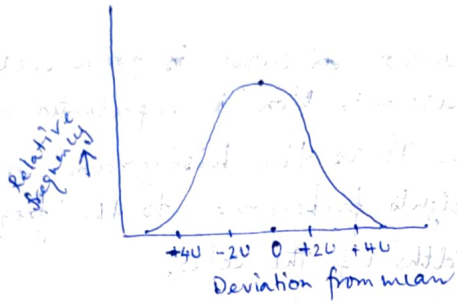
Constructing error curve:

Small undetectable uncertainties produces a detectable random error in according to the following way. For the sake of simplicity if we consider a situation, where first four small random errors combine to give an overall error (ii) that each error has an equal probability of occuring and that each can cause the final result to be high or low by a fixed amount $\pm U$.

Then that all the possible ways that four errors can combine to give the deviations from the mean are

Combination of uncertainties	Magnitude of random error	No. of combinations	Relative frequency
$+U_1 + U_2 + U_3 + U_4$	$+4U$	1	$\frac{1}{16} = 0.0625$
$-U_1 + U_2 + U_3 + U_4$ $U_1 - U_2 + U_3 + U_4$ $U_1 + U_2 - U_3 + U_4$ $U_1 + U_2 + U_3 - U_4$	$+2U$	4	$\frac{4}{16} = 0.250$
$-U_1 - U_2 + U_3 + U_4$ $+U_1 + U_2 - U_3 - U_4$ $+U_1 - U_2 + U_3 - U_4$ $-U_1 + U_2 - U_3 + U_4$ $-U_1 + U_2 + U_3 - U_4$ $U_1 - U_2 - U_3 + U_4$	$0U$	6	$\frac{6}{16} = 0.375$
$+U_1 - U_2 - U_3 - U_4$ $-U_1 + U_2 - U_3 - U_4$ $-U_1 - U_2 + U_3 - U_4$ $-U_1 - U_2 - U_3 + U_4$	$-2U$	4	$\frac{4}{16} = 0.250$
$-U_1 - U_2 - U_3 - U_4$	$-4U$	1	$\frac{1}{16} = 0.0625$

These indicate that only one combination leads to a deviation of +4U, four combinations give a deviation of +2U, and six give a deviation of 0U. The negative errors have the same relationship. This ratio of 1:4:6:4:1 is a measure of the probability for a deviation of each magnitude. When the same procedure is applied for a large no. of measurements, and plotted, the frequency distribution of bell-shaped normal error curve or Gaussian curve is obtained.



Properties of a normal error curve:

A normal error curve has several general properties:-

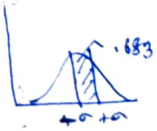
(a) The mean occurs at the central point of maximum frequency.

(b) There is a symmetrical distribution of positive and negative deviations about the maxima.

(c) There is a exponential decrease in frequency as the magnitude of the deviation increases. Thus, small uncertainties are observed much more often than very large ones.

Significance of area under a normal error Curve:

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$



Area within $\pm 1\sigma$; area $= \int_{-\sigma}^{+\sigma} \frac{e^{+(x-\mu)^2/2\sigma^2}}{\sigma\sqrt{2\pi}} dx$

If we consider a new variable z as $z = \frac{x-\mu}{\sigma}$

The integration becomes, area $= \int_{-1}^{+1} \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} dz$

$= 0.683$

Thus, 68.3% of the area beneath a Gaussian curve for a population lies within one standard deviation ($\pm 1\sigma$) of the mean μ It indicate that there are 68.3 chances in 100 that the random uncertainties of any single measurement is no more than $\pm 1\sigma$.

Similarly, it can be shown that approximately 95.4% of all data values are within $\pm 2\sigma$ of the mean and 99.7% are within $\pm 3\sigma$, and so on.

Histogram:

When the normal error curve is for a set of measurements in a particular determination is represented as a bar plot, it is known as histogram. Thus the histogram consists of contiguous columns of hlights proportional to the frequencies, erected upon the full widths of the cells.

Ex:- The Histogram for a set of 60 measurements in the absorbance value may be utilised as follows to construct a representative histogram —

Measurements are: 0.458, 0.450, 0.465, 0.452, 0.452, 0.447, 0.459, 0.451, 0.446, 0.467, 0.452, 0.463, 0.456, 0.456, 0.489, 0.454, 0.456, 0.441, 0.457, 0.459, 0.462, 0.450, 0.454, 0.446, 0.464, 0.461, 0.463, 0.457, 0.460, 0.451, 0.456, 0.455, 0.451, 0.462, 0.451, 0.469, 0.458, 0.458, 0.456, 0.454, 0.450, 0.455, 0.456, 0.456, 0.459, 0.454, 0.455, 0.458, 0.457, 0.456, 0.455, 0.460, 0.456, 0.463, 0.457, 0.456, 0.457, 0.453, 0.455, 0.453

First, the results are arranged in order from lowest to highest.

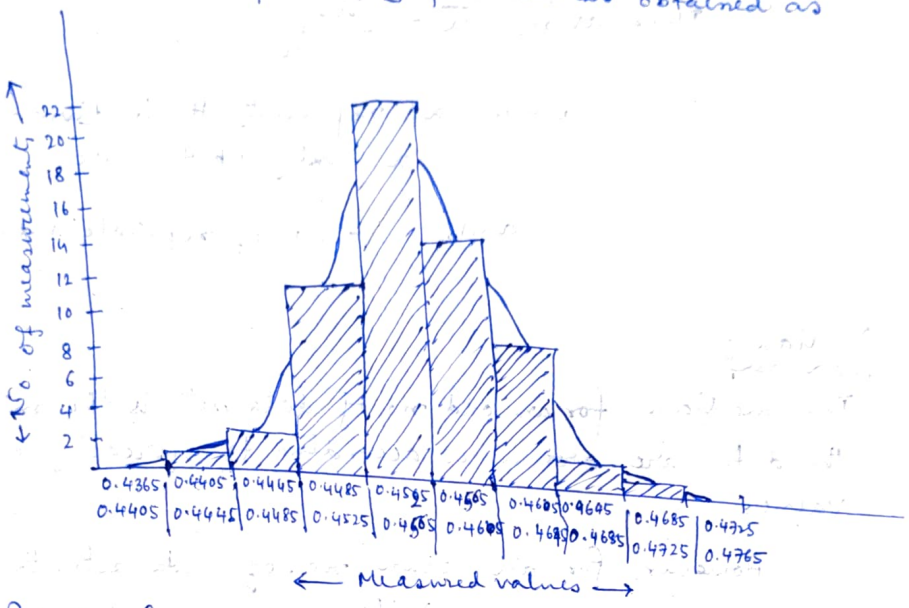
Second, The data are condensed by grouping them into cells. It is done by deviding the range into a convanient no. of intervals and then no. of values are counted which falls within each cell. Generally, ranges are equally devided into equal intervals to get the midpoint and boundaries.

For the above vals, the groupings are as follows.

Cell mid point	Cell boundaries	No. of values
	0.4405	
0.4425	0.4445	1
0.4465	0.4485	3
0.4505	0.4525	11
0.4545	0.4565	21
0.4585	0.4605	14
0.4625	0.4645	7

Cell midpoint	Cell boundaries	No. of values
0.4665	0.4685	2
0.4705	0.4725	1

When the no. of measurements are plotted against measured values, Therefore, histograms is pictorially presented as obtained as



However, Relative frequency curve is constructed by plotting frequencies at cell mid points and connecting the points with straight lines.

Mistaxes

Mistaxes are crude errors which greatly distorts the analytical results. They include errors caused by incorrect counting of weights or wrong readings on the scale in weighing, wrong burette reading in volumetric analysis, errors caused by spilling of the solutions or precipitate during the determination, etc.

A mistaxe maxes the result of the given determination incorrect, and it is therefore rejected when the average of a series of replicate determinations is found.