**Evaluation of Analytical Data**

Any quantitative analysis deals with estimations and reliability and corrections of result are utmost important. This is achieved by repeating the experiment several times and then applying statistical procedures to the analytical results.

**1. True, Standard and Observed value**

The result obtained in any experiment is an observed value. In order to determine its corrections it is essential to know the true or correct or the actual value. True values can only be determined if a direct count is made or a quantity is assigned a specific value viz C has atomic weight of 12.000. In a few cases it is not possible to know the true values of a quantity. However an efficient analyst using the most suitable method and good quality apparatus may arrive at a result which can be regarded as true value but it is the called standard value.

**2 . Errors**

The difference between a true or standard value and the observed value is an error. A quantitative estimation not only is collection of data but also involves, understanding the principles involved and possible sources of errors.

***Absolute error* :** It is the difference between the observed or experimental value and the true or standard value. Example- Standard value of gold in a given sample is 99% and the analyst determines a value of 98.2%, the error of determination is (99-98.2) *= 0.8%.*

***Relative error :*** It is obtained by dividing absolute error with the true value. It is expressed in percent or parts per thousands .

In the above example, the relative error is

 $\frac{0.8}{99}$ x 100 = 0.8080

or $\frac{0.8}{99}$ x 1000 = 8.8 ppt (parts per thousand)

**Classification of errors :** In order to keep the error as small as possible, it is necessary to know the types of errors and their possible causes. Errors are broadly classified as :-

(I) Determinate or systematic errors

(II) In determinate or Random errors

**(1) Systematic (determinate) errors –** These are the errors where the cause is known and can be avoided or whose magnitude can be measured. The most important of them are.

1. **Operational and Personal errors :** These are due to the factors for which individual analyst is responsible. These are mostly physical in nature and occur when proper analytical techniques are not followed. Examples:
	1. Mechanical loss of substance in course of analysis
	2. Under washing or over washing
	3. Ignition at incorrect temperatures
	4. Insufficient cooling of precipitation
	5. Allowing exposure of hygroscopic material etc.

Personal errors arise out of constitutional inability to take accurate observation., viz colour blindness, double vision, inability of observe sharp colour charge etc.

1. **Instrumental or Reagent errors** - these errors occur mostly due to
	1. Faulty balances
	2. Uncalibrated weights and glassware
	3. Reaction between reagents and containers
	4. Changes at high temperature
	5. Impure reagents
2. **Errors of method -** These arise out of procedural condition such as -
	1. Incorrect sampling and incompleteness of reaction
	2. Co-precipitation and post precipitation
	3. Occurrence of side reactions
	4. Titration error

1. **Additive errors** – sometimes the value of error is constant in a series of determination and is independent of the amount of sample taken for analysis. Example- In a titration 0.1 ml of titrant is added to see a sharp colour change. If the true value is 10.0 ml the observed value will be 10.1 ml. Now, if the sample is doubled and the standard value is 20.0 ml, the observed value will be 20.1 ml; error remains 0.1 irrespective of sample site.
2. **Proportional errors :** Magnitude of error depends on the sample size. It may not vary uniformly with the amount of sample but may show an increase in error with increase in sample size.

**Minimisation of errors –** The following methods can minimize **determinate errors.**

1. Calibration of the apparalus - All the glassware and weights should be calibrated and corrections applied.
2. Quality reagents - High quality reagents with known and very small impurity should be used.
3. Running a blank titration - this consists of carrying out a determination under identical conditions without the sample. The aim is to find out the effect of impurities introduced through the reagents and vessels.
4. Running a controlled determination – A parallel determination is carried out by using nearly the same amount of constituent that is present in the sample under identical conditions. The weight of the constituent in the sample is calculated using the following relationship.

$\frac{Result of known sample}{Result of unknown sample }$ = $\frac{Weight of constituent in known sample }{Weight of constitutent in unknown sample }$

1. Use of independent method - A determination is carried out by an entirely different method and the two results are compared. Example- Nickel can be estimated gravimetrically using dimethyl glyoxime and titrimetrically using EDTA.
2. Standard addition – A small known amount of constituent being determined is added to the sample, which is then analysed for total amount. The difference between the results for samples with and without added constituent indicates the recovery of the added amount. A satisfactory recovery establishes the accuracy of results.
3. Isotopic dilution - A known amount of the element being determined containing an radioactive isotope is mixed with the sample and the element is estimated. The resultant radioactivity indicates amount of sample.

**(II) Random (Indeterminate) error :** These occur due to causes over which analyst has no control. They are intangible and incapable of analysis and the cause may or has no knowledge or control. If a sufficiently large number of readings are taken then these errors can be interpreted by applying statistical tools.

**3. Accuracy**

The accuracy of a determination may be defined as the concordance between the observed value and true or standard value. It can be found out by absolute and comparative method.

1. **Absolute method :** The sample is synthesized using known quantities of the constituents and a primary standard is obtained. The true values of amount of constituents are hence known. Now the sample is analysed by some method and an observed value is obtained.
2. **Comparative method :** In some cases, it is not possible to prepare the synthetic compound, viz ores. Analytical data from standard agencies are obtained, results compared and the error is ascertained. This method involves secondary standards.

**4. Precision**

It refers to the similarity of a series of measurements of the same quantity. If an experiment is done may times, each time there may be some variation. Precision indicates nearness between various measurements.

**Accuracy expresses the correctness of a measurement whereas precision expresses the reproducibility of results.** Precision always accompanies accuracy but high degree of precision does not imply accuracy. Example- Two analysts X and Y determined the % of gold in the same coin. The standard value is 100%. Their results were as follows.

 Analyst A – 100, 99.8, 99.6 , 99.7, 99.2

 Average value – 99.66 error – 0.34 %

 Analyst B -98.80, 98.82, 98.86, 98.82, 98.81

 Average value – 98.82 error – 1.18 %

 The result indicate that

1. The values obtained by analyst A are accurate, very close to true value but the precision is poor as there is variation in results.
2. The values of analyst B are very precise but are less accurate.

**5. Mean and Standard Deviation**

The minimize the errors, a determination is repeated number of times and values obtained are examined statistically to get the final result.

**Average or mean** is a measure of central tendency i.e. other values are around the average values. It is the arithmetic mean of different values.

If there are ‘n’ observations arranged in the increasing order.

 x1, x2,x3 ………….Xn-1,xn

Arithmetic mean = $\frac{Sum of different values }{Number of values }$ = $\frac{x1+x2+….+xn}{n }$ = $x$

**Deviation** (d), is the difference between the measured value and the average value, i.e. x1, - x2, =d1 ; d1 is the deviation in the first determination.

d= $\frac{d1+d2+d3 ….+dn}{n }$

The (+) or (-) sign of individual deviation is disregarded;

**Relative mean deviation** is given by-

$\frac{Mean deviation}{Amount of the constituent }$ x 100

**The standard deviation (root mean square deviation)** is the square root of the mean of the squares of the differences between the values and the mean of those values.



The square of the standard deviation (s) is called **variance**. Another measure of precision , known as **Relative Standard Deviation** (R.S.D) is given by :

 R.S.D. = $\frac{S}{x}$

R.S.D expressed as percentage is called the coefficient of Variation (C.V.) :

 C.V. = $\frac{Sx100}{x}$

Example – Analyses of a sample of copper are gave the following % value for copper content, 8.02, 8.11, 8.07, 8.06, 8.09, 8.10, 8.11,8.02, 8.08, 8.07. Calculate the mean, standard deviation and C.V.

|  |  |  |
| --- | --- | --- |
| Result (x)  | x – $x $ | (x –x)2  |
| 8.028.118.078.068.098.108.118.028.088.07 | -0.050.040.020.00-0.010.030.01-0.05 0.010.02 | 0.00250.00160.00040.00000.00010.00090.00010.00250.00010.0004 |
| $∑$ x = 80.73 |  | $∑$ (x-x)2 = 0.0076 |

Mean X = $\frac{∑x}{n}$ = $\frac{8073}{10}$ = 8.073

 S = $\sqrt{\begin{array}{c}0.0076\\\\_\\_\\_\\_\\_\\_\\9\end{array}}$

 = $\sqrt{0.00084}$

S= $\pm $ 0.0289 percent

C.V. = $\frac{0.0289 x100}{8.07 }$ = 0.358 percent

 **6. Distribution of random errors**

If a large number of replicate readings, at least 50, are taken, the result will be distributed about the mean value in a symmetrical manners. The mathematical model that best satisfies the distribution of random errors is called Normal (Gaussain) distribution. It is a bell shaped curve that is symmetrical about the mean value (Fig .1)

 M -II M -I M M+I M+II

Fig. 1 Normal (Gaussian) distribution of random errors

In this type of distribution 68 percent of all values will fall within one standard deviation on either side of the mean, 95 percent will fall within two standard deviation and 99. Percent within three standard deviation. For the worked example (section 4.5) the standard deviation is $\pm $ 0.0289 percent. If results are normally distributed then 68% (approximately 7 out of 10 results) will be between $\pm $ 0.0289, X 2 = $\pm $ 0.0578 percent (allowing two standard deviations) (i.e. 9 results are between 8.13 -8.01). Only 5% probability is therefore a result to differ from mean by more than $\pm $ 0.0578%.

 **7. Confidence Interval (limit)**

When a small number of observation is made, the value of standard deviation(s), does not reveal as to how close the sample mean x is to true mean. It is, however, possible to calculate a confidence interval to estimate the range within which true mean may be found. The limits of this confidence interval, called confidence limit ($μ$) are given us :

Confidence limit ($μ$), for n replicate measurements = x $\pm $ = $\frac{ts}{n}$

Where t is a parameter that depends upon the number of degrees of freedom d the confidence level required.

Example - The mean (x) of four determination of nickel in an alloy is 9.35 percent with a standard deviation of 0.15 percent. Calculate the 95% confidence limit for the given value.

From the t-tables, the value of t for 95% confidence level with (n-1) i.e. three degrees of freedom is 3.18. So confidence level $μ$ is

$μ$ = 9.35 $\pm $ $\frac{3.18x0.15}{4}$

= 9.35 $\pm $ 0.238 percent

Thus, there is 95% confidence that true value of content of the alloy lies in the range of nickel 9.59 to 9.11 percent.

**Questions**

1. What is an Error ? Define absolute and relative error .
2. Discuss systematic and random errors. What are their causes ?
3. Write a note on minimisation of errors
4. Define Accuracy and Precision.
5. What is a) Mean deviation b) Standard deviation c) Relative standard deviation d) Confidence limit.
6. Explain Normal distribution curve of random errors.
7. Write a note on Significant Figures and their importance.