Chapter (2)

Errors and their Analysis

Errors

It is the discrepancy between the results of the measurement and the value of the quantity to be measured.

Sources of Errors

- This dispersion between the measured and true values of a certain quantity has origin in several components of the measuring system.

1. Within operator variation.
2. Between operators variation.
3. Material variations.
4. Test equipment variation.
5. Test procedure variation.
6. Between laboratory variation.
7. Noise.
8. Response Time.
9. Design Limitation.
10. Resolving Power.
11. Energy Exchange by interaction.
12. Observation and Interpretation.
13. Deterioration of measuring instrument.

Types of Errors

1. **Controllable Error**. Those are the errors that should not occur, however, may be corrected and eliminated by more careful work and attention.
2. **Random Error**. Those are types of errors that are inherited in the measurement process or instrument.

Types of Controllable Errors

The errors in a scale and pointer type of measuring instrument can be of following three types:

1) **Assembly errors**. These can be due to the following :

- **Displaced scale (incorrect fitting of the scale zero with respect to the actual zero position of the movement).**
- **Non-uniform division.**
- **Bent or distorted pointer**.

Errors of this type can be easily discovered and rectified as they remain constant with time.

2) **Systematic Errors**. **Those are types of errors that *in the course of a number of measurements, made under the same conditions, of the same value of quantity, they either remain constant in absolute value and sign or vary according to some definite law*.**

- **Constant Systematic Errors (constant in nature).**
- **Environmental Errors (Caused by changing the environmental conditions from the calibrated ones).**

How to correct Environmental Errors?

1. Using instrument in controlled conditions of pressure, temperature and humidity in which it was originally assembled and calibrated.
2. If above is not possible then deviations in local conditions must be measured and suitable corrections to instrument readings applied.
3. Automatic compensation using sophisticated devices for such deviations is also possible and usually applied.
4. Altogether new calibration may be made in the new conditions.

The usual environmental conditions are :

Ambient Temperature = 20 °C, Barometric Pressure = 76 mm Hg, and Humidity =10 mm Hg

Let;

T1 = Temperature of the work piece (°C)

T2 = Temperature of the measuring equipment (°C)

α 1 = Coefficient of thermal expansion of the work piece (1/ °C)

α 2 = Coefficient of thermal expansion of the measuring equipment (1/ °C)

L_2 = length to be measured (m)

L_T = True length of dimension at 20 °C

$$L_{True} = L_2 * [1-(\Delta \alpha * \Delta T)]$$

$$\Delta \alpha = (\alpha_1 - \alpha_2), \ \Delta T = (T_2 - 20)$$

(3) **Instrumental Errors**.

- Friction Errors.
- Scale Errors.
- Zero Errors.

(4) **Catastrophic Errors** (Errors of large magnitude in nature).

- Misreading Errors (either due to misreading or error in display).
- Arithmetic Errors.

(5) **Parasitic Errors** (Errors which result from an incorrect execution of the measurement process).

- Error of Method.
- Observation Error.
- Parallax Error (When there is a big gap between the scale and pointer).
- Misalignment Error.
- Interpolation Error (due to inexact evaluation of the position of the index with reference to 2 adjoining points).

6) **Mechanical Vibration.**

7) **Backlash in movement.**

8) **Finite dimensions of the pointer and scale divisions.**

9) **Hysteresis in elastic members or electric machines.**

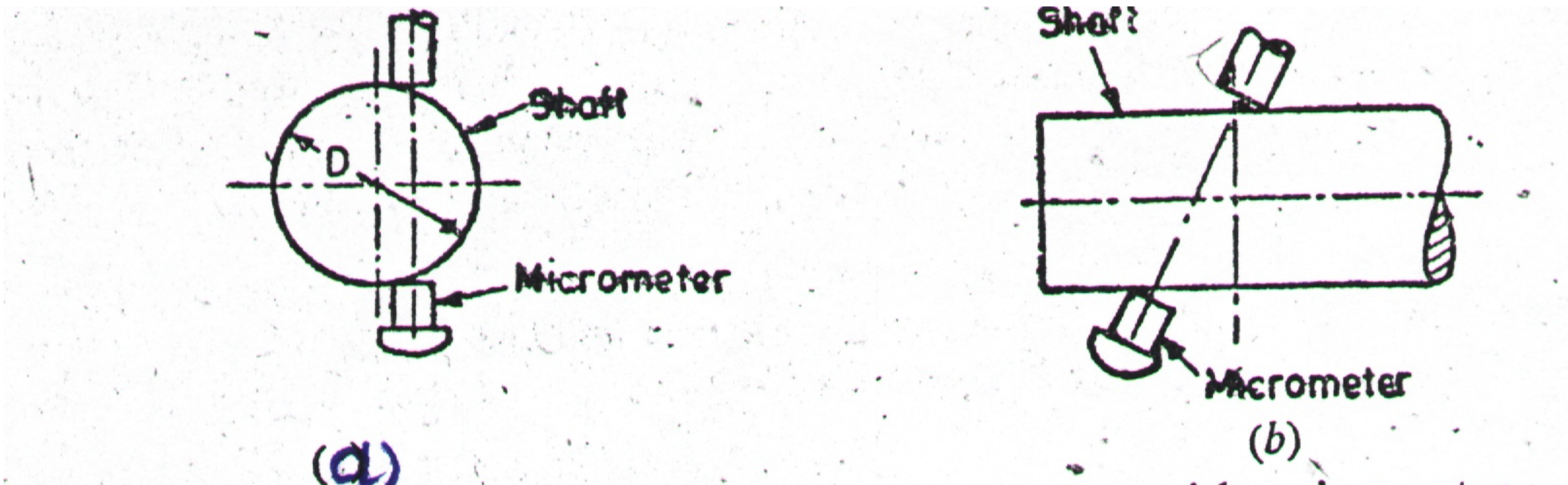

Fig. 7·3. Wrong ways of measurement of diameter with micrometer.

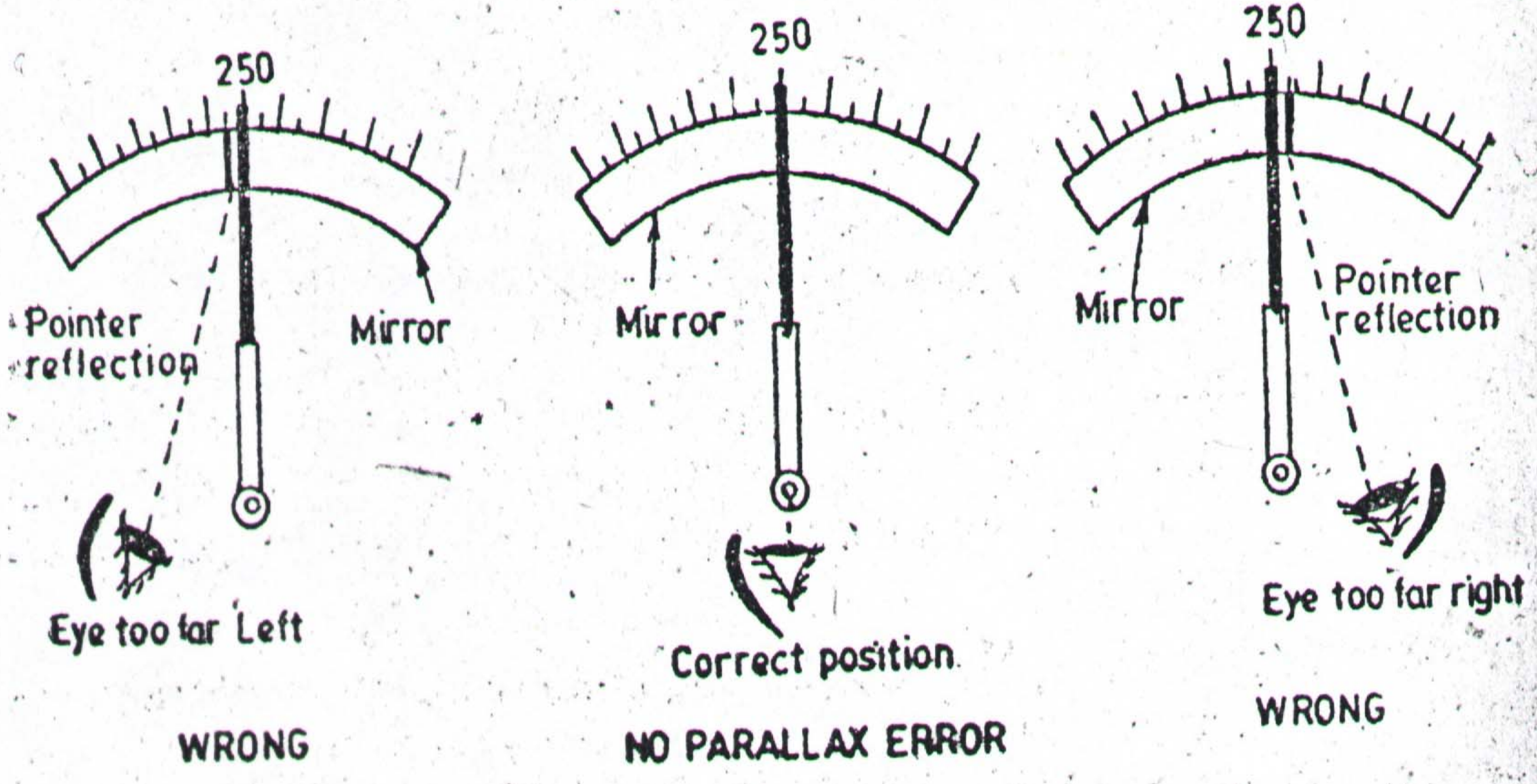

Fig. 7·6. Errors due to Parallax.

nt of pointer and the

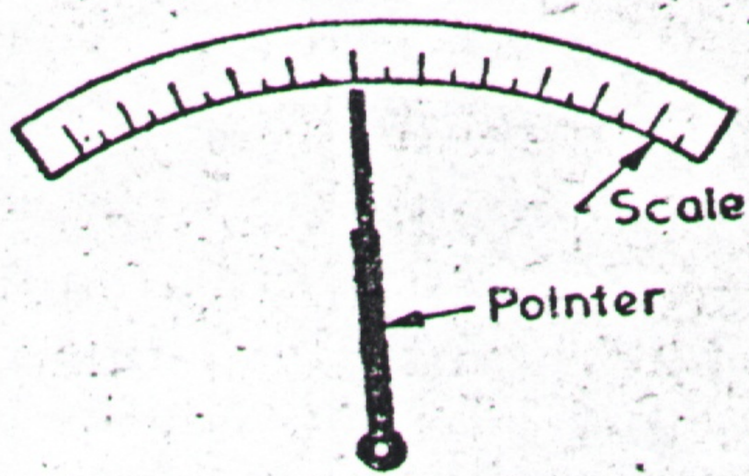

Fig. 7 7. Arrangement showing scale and pointer in the same plane.

Random Errors

An error which varies in an unpredictable manner in absolute value and in sign when a large number of measurements of the same value of a quantity are made under effectively identical conditions.

Statistical Treatment of Data

The experimental data is obtained in two forms of tests :

(i) ***Multisample Test***. In this test, repeated measurement of a given quantity are done using different test conditions such as employing different instruments, different ways of measurement and by employing different observers.

(ii) ***Single Sample Test***. A single measurement (or succession of measurements) done under identical conditions excepting foretime is known as single-sample test.

- In order to get the exact value of the quantity under measurement, tests should be done using as many different procedures, techniques and experimenters as practicable.

Histogram

- This histogram of Figure represents these data where the ordinates indicate the number of observed reading's frequency of occurrence of a particular value. A histogram is also called a **frequency distribution curve**.

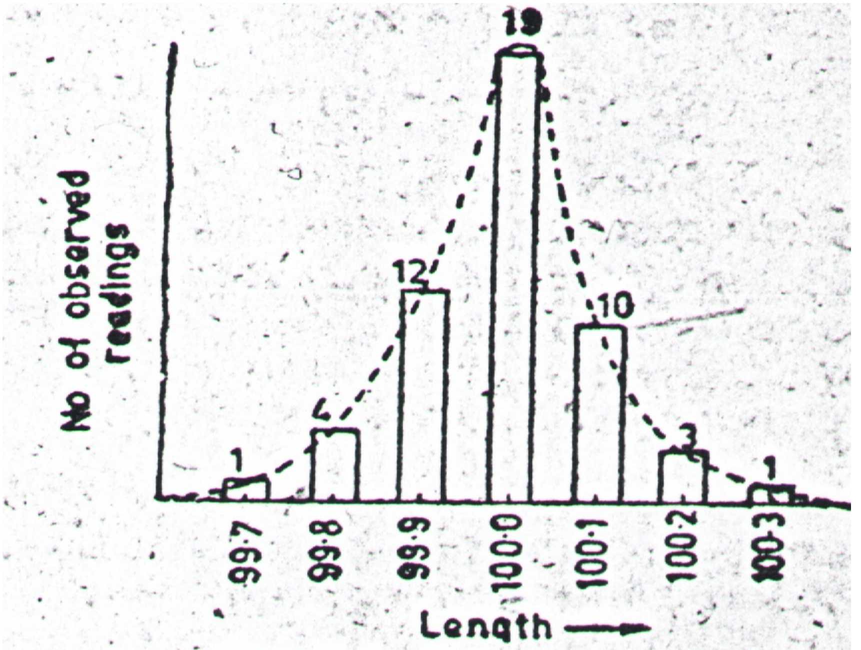

Fig. 7.8. Histogram.

■ Arithmetic Mean ($\bar{x}$)

The most probable value of measured variable (variate).

[redacted]

F = frequency of occurrence of “x” value

Dispersion from the Mean

The property which denotes the extent to which the values are dispersed about the central value.

- Other names used are **spread** or **scatter**.
- In one case (curve 1) the values vary from x1 to x2 and in other case (curve 2) the values vary from x3 to x4, though their central value is the same.
- Clearly set of data represented by curve 1 has a smaller dispersion, than that of the data represented by curve 2
- It is an indication of the **degree of consistency (precision)** and regularity of the data.

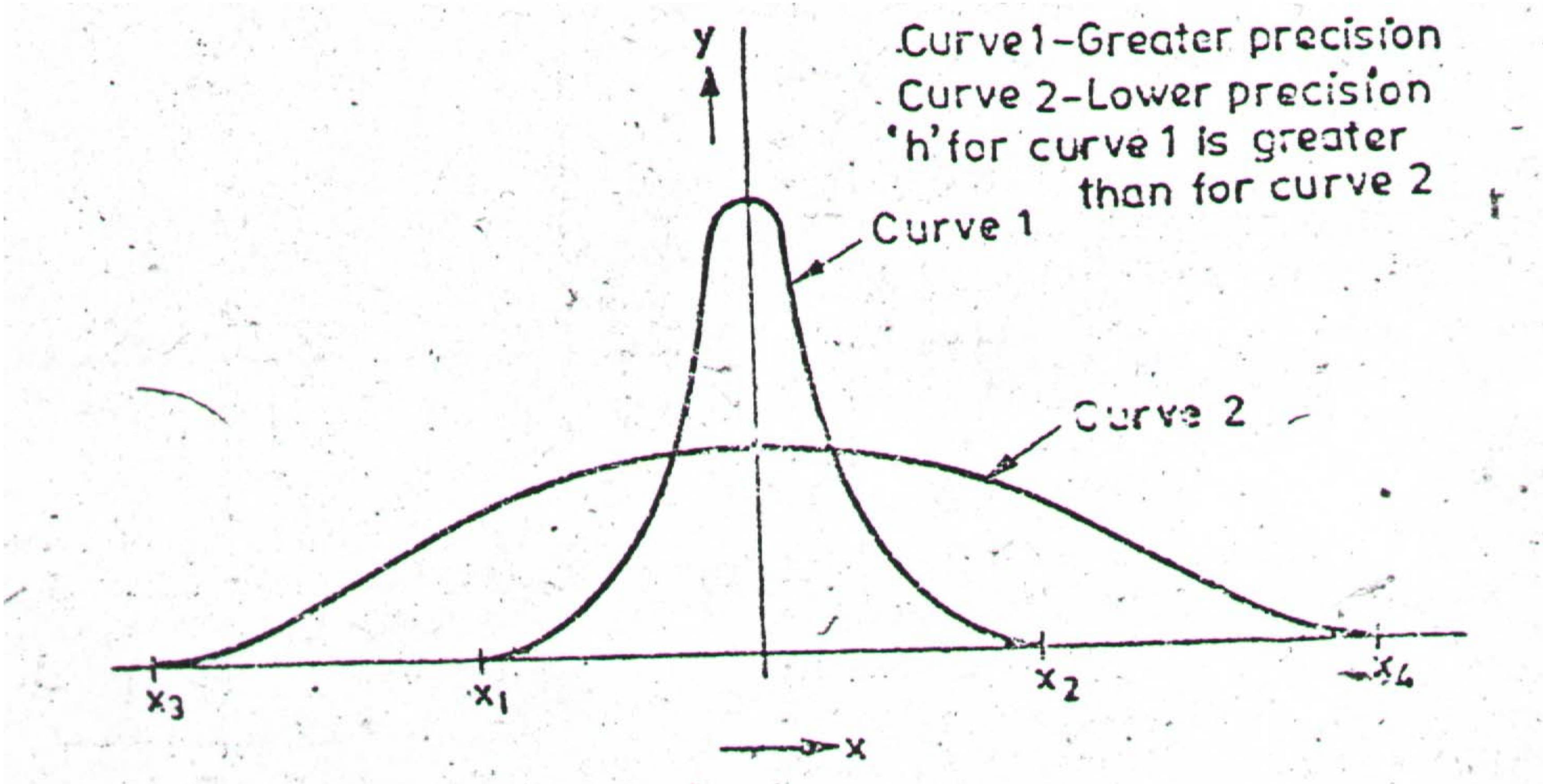
Curve 1–Greater precision
Curve 2–Lower precision
'h' for curve 1 is greater than for curve 2
Curve 1
Curve 2
y
x
x_3
x_1
x_2
x_4

Fig. 7·9.

- A large dispersion indicates that some factors involved in the measurement process **are not under close control** and therefore it becomes difficult to estimate the measured quantity with confidence; and definiteness.
- There are certain terms which must be defined as they form the basis of defining the measure of dispersion of data, these are :

Range.

Is the difference between greatest and least values of data.

For example in Figure 7.8 the range of curve 1 is $(x_2 – x_1)$ and that of curve 2 is $(x_4 – x_3)$.

▪ Deviation (d).

Is the departure of the observed reading from the arithmetic mean of the group of readings. Let the deviation of reading x_1 be d_1 and that of reading x_2 be d_2, ………..etc. Then

$$d_i = x_i - x$$

▪ Average Deviation (D) .

It *is defined as the sum of the absolute values of deviations divided by the number of readings.*

- The average deviation is an indication of the precision of the instruments used in making the measurements.
- Highly precise instruments yield a low average deviation between readings.

$\overline{D} = \frac{\sum |d.f|}{\sum f}$

The average deviation is given by :

$$D = \sum | d*f | / \sum f$$

- **Standard Deviation (σ).**

Is the square root of the sum of the individual deviations squared, divided by the number of readings. It tells how tightly the data are clustered around the average value. It also gives the expected variation of the data around the mean.

$$\sigma = (\sum d*f^2 / \sum f)^{0.5}$$

- **Variance.**

It measures how spread out the data is. The variance is the mean square deviation, which is the same as S.D. except that square root is not extracted.

$$V = \sigma^2$$

Precision Index (Y).

This is given by the following equation :

$$Y = (h/\sqrt{\pi}) \int \exp(-h^2x^2)\, dx$$

when x=0, we have :

$$Y = (h/\sqrt{\pi})$$

- The larger the value of 'h', the sharper the curve.
- A sharp curve evidently indicates that the deviations are more closely grouped together around deviation x=0.
- It is clear that the probability that a variate lies in a given range becomes less as the deviation of the range becomes greater.
- A large value of 'h' represents high precision of the data because the probability of occurrence of variates in a given range falls off rapidly as the deviation increases because the variates tend to cluster (become closer) into a narrow range.

- On the other hand, a small value of ‘h’ represents low precision because the probability of occurrences of variates in a given range falls off gradually as the deviation increases ; this is because the variates are spread over a wide range.

The value of (h) is given by :

$$\mathbf{h = (\sigma\sqrt{2})^{-1}}$$

Probable Error of One Reading (r).

In order to check the validity of the selection of (X) as the true value probable error is used as follows.

A convenient measure of precision is the quantity ‘r’. It is called Probable Error or simply P.E. The reason for this name is the fact mentioned above that half the observed values lie between the limits $\pm$r. If we determine r as the result of n measurements and then make an additional measurement, the chances are 50 -50 % that the new value will lie between -r and +r. That is, the chances are even that any one reading will have an error no greater than $\pm$r.

The location of point r can be found by the following:

$$0.5 = (h/\sqrt{\pi}) \int \exp(-h^2x^2)\, dx$$

This gives

$$r = 0.4769 / h$$

■Standard Deviation of the Mean (σ_m)

When we have a multiple sample data, it is evident that the mean of various sets of data can be analyzed by statistical means. This is done by taking standard deviation of the mean. The standard deviation of the mean is given by:

$$\sigma_m = \sigma / (\sum f)^{0.5}$$

■Standard Deviation of the S.D. (σ_σ)

For a multiple sample data, the standard deviation of the standard deviation is:

$$\sigma_\sigma = \sigma_m / (2)^{0.5}$$

Example (2)

In a test, temperature is measured 100 times with variations in apparatus and procedures. After applying the known corrections, the results are shown in Table below.

Calculate:

(a) arithmetic mean, (b) mean deviation, (c) standard deviation, (d) the probable error of one reading, (e) the standard deviation and the probable error of the mean, (f) the standard deviation of the standard deviation.

T °C	397	398	399	400	401	402	403	404	405
F	1	3	12	23	37	16	4	2	2

T °C	F	T*F	d=(T-Tavg)	abs(d*F)	d^2	d^2*F
397.000	1	397	-3.780	3.78	14.2884	14.2884
398.000	3	1194	-2.780	8.34	7.7284	23.1852
399.000	12	4788	-1.780	21.36	3.1684	38.0208
400.000	23	9200	-0.780	17.94	0.6084	13.9932
401.000	37	14837	0.220	8.14	0.0484	1.7908
402.000	16	6432	1.220	19.52	1.4884	23.8144
403.000	4	1612	2.220	8.88	4.9284	19.7136
404.000	2	808	3.220	6.44	10.3684	20.7368
405.000	2	810	4.220	8.44	17.8084	35.6168
Sum	100	40078		102.84		191.16

Mean Temperature =	400.78 ^{o}C
Mean Deviation =	1.0284
Standard Deviation =	1.38260623
Variance =	1.9116
Probable Error of one reading (r1) =	0.93256791
Probable Error of the mean (rm) =	0.09325679
Standard deviation of the Mean =	0.13826062
Standard deviation of the standard deviation =	0.09776502

Specifying Measurement Data

- After doing the statistical analysis of the multisample data, we must specify the results. The results are expressed as deviations about a mean value. The deviations are expressed as:

(i) <u>Standard deviation</u> ($\pm\sigma$): The result is expressed as:

$$X \pm \sigma$$

- The error limit in this case is the standard deviation. This means that 0.6828 (about 68%) of the readings are within the limits $\sigma = \pm 1$ and the odds are 2.15 to 1. Thus there is approximately a 2 to 1 possibility that a new observation will fall beyond this limit.

(ii) Probable error ($\pm$ r): The result is expressed as:

$$X \pm 0.6745\ \sigma$$

- This means that 50% of the readings lie within this limit and the odds are 1 to 1. This means that there is an even possibility that a new reading will lie within these limits.

(iii) $\pm 2\sigma$ limits: In case we want to increase our probability range we specify the results as:

$$X \pm 2\ \sigma$$

- Thus we assume that 0.9546 (or about 95%) of readings fall within these limits. These odds in this case are 21 to 1.

(iv) $\pm 3\sigma$ limits: The results in this case are expressed as:

$$X \pm 3\sigma$$

- The maximum or boundary error limit is $\pm 3\sigma$. The probabi-lity in this case is 0.9974. This means that 99.74% of the obser-vations will fall within this limit. In other words we can say that there is a possibility of only 26 readings out of 1000 to fall beyond these limits. Thus practically all the observations are included in this limit. The odds of any observation falling Out of this limit are 256 to 1.

Criteria for Data Rejection

- In most of the experiments, the experimenter finds that some of the data points are noticeably different from the majority of the data. If these data points were obtained under abnormal conditions involving gross blunders and the experimenter is sure about their dubious nature they can be discarded straight away.
- The experimenter cannot reject a data simply because it is diff-erent from the others. He must rely on, certain standard mathema-tical methods for rejecting any experimental data. There are many methods available for assessing whether the data be rejected or retained These methods are discussed.

Chauvenet's Criterion

- Suppose "n" observations arc made for measurement of a qua-ntity. We assume that "n" is large enough that the results will follow a normal Gaussian distribution. This distribution may be used to compute the probability that a given reading will deviate by a certain amount from the mean.
- ***Chauvenest's criterion specifies that*** a reading may be rejected if the probability of obtaining the particular deviation from the mean is less than 0.5n.
- Table 7.6 below gives the values of the ratio of deviation to standard deviation for various values of "n" according to this criterion.

- Table 7.6. Chauvenest's Criteria for Rejection of Data.

Number of Readings	Ratio of maximum acceptable deviation to standard deviation
2	1.15
3	1.38
4	1.54
5	1.65
6	1.73
7	1.80
10	1.96
15	2.13
25	2.33
50	2.57
100	2.81
300	3.14
500	3.29
1000	3.48

- When applying Chauvinist’s criterion in order to eliminate any dubious data, follow these steps :

1. The mean value and the standard deviation are first calculated using all data points.
2. The deviations of individual readings are then compared with standard deviation.
3. If the ratio of deviation of reading to the standard deviation exceeds the limits given in Table 7.6, that reading is rejected.
4. The mean value and the standard deviations are again calculated by excluding the rejected reading from the data.

Example (3)

- Solve Example (2) by finding the data to be rejected (if any).

T	F	T*F	d=(T-Tavg)	abs(d*F)	d^2	d^2*F	SD/SDmax
397.00	1	397	-3.780	3.78	14.2884	14.2884	2.73396713
398.00	3	1194	-2.780	8.34	7.7284	23.1852	2.01069540
399.00	12	4788	-1.780	21.36	3.1684	38.0208	1.28742367
400.00	23	9200	-0.780	17.94	0.6084	13.9932	0.56415194
401.00	37	14837	0.220	8.14	0.0484	1.7908	0.15911978
402.00	16	6432	1.220	19.52	1.4884	23.8144	0.88239150
403.00	4	1612	2.220	8.88	4.9284	19.7136	1.60566323
404.00	2	808	3.220	6.44	10.3684	20.7368	2.32893496
405.00	2	810	4.220	8.44	17.8084	35.6168	3.05220669
Sum	100	40078		102.84		191.16	

- As noticed, the deviation ratio for the final reading exceeded the 2.81 specified by Chauvenest's criteria, hence all the 405^{o}C readings must to be removed and new calculations of data be done as shown below.
- Now, for the recalculations, the total number of readings will reduce to 98. Based on Chauvenest's criteria, the maximum deviation ratio should not exceed 2.57.
- The Table below shows the second cycle of calculations.

T	F	T*F	d=(T-Tavg)	abs(d*F)	d^2	d^2*F	SD/SDmax
[redacted]							
398.00	3	1194	-2.780	8.34	7.7284	23.1852	2.010695
399.00	12	4788	-1.780	21.36	3.1684	38.0208	1.287423
400.00	23	9200	-0.780	17.94	0.6084	13.9932	0.564151
401.00	37	14837	0.220	8.14	0.0484	1.7908	0.159119
402.00	16	6432	1.220	19.52	1.4884	23.8144	0.882391
403.00	4	1612	2.220	8.88	4.9284	19.7136	1.605663
404.00	2	808	3.220	6.44	10.3684	20.7368	2.328934
Sum	98	39268		94.4		155.5432	

- Again, the deviation ratio for the first reading exceeded the 2.57 specified by Chauvenest's criteria, hence all the 405oC readings must to be removed and new calculations of data be done as shown below.
- Now, for the recalculations, the total number of readings will reduce to 97. Based on Chauvenest's criteria, the maximum deviation ratio should not exceed 2.57.
- The Table below shows the third cycle of calculations.

T	F	T*F	d=(T-Tavg)	abs(d*F)	d^2	d^2*F	SD/SDmax
398.00	3	1194	-2.780	8.34	7.7284	23.1852	2.010695403
399.00	12	4788	-1.780	21.36	3.1684	38.0208	1.287423675
400.00	23	9200	-0.780	17.94	0.6084	13.9932	0.564151948
401.00	37	14837	0.220	8.14	0.0484	1.7908	0.15911978
402.00	16	6432	1.220	19.52	1.4884	23.8144	0.882391508
403.00	4	1612	2.220	8.88	4.9284	19.7136	1.605663235
404.00	2	808	3.220	6.44	10.3684	20.7368	2.328934963
Sum	97	38871		90.62		141.2548	

Limiting Error (δA)

- The accuracy and precision of an instrument depends upon :

1. Its design,
2. The material used , and,
3. The workmanship that goes into making the instrument.

- The choice of an instrument for a particular application depends upon the accuracy desired. If only a fair degree of accuracy is desired, it is not economical to use expensive materials and skill into the manufacture of the instrument. But an instrument used for an application requiring a high degree of accuracy has to use expensive material, and a highly skilled workmanship.

- The economical production of any instrument requires the proper choice of material, design and skill. In order to assure the purchaser of the quality to the instrument, the manufacturer guarantees a certain accuracy.
- In most instruments the accuracy is guaranteed to be within a certain percentage of full scale reading. Components are guaranteed to be within a certain percentage of the rated value. Thus the manufacturer has to specify the deviations from the specified value of a particular quantity. *The limits of these deviations from the specified value* are defined as ***Limitation errors or Guarantee errors***.

- The magnitude of a quantity having a specified magnitude A1 and a maximum error or limiting error of ± δA must have a-magnitude between the limits A1- δA and A1+δA or

$$A = A1 \pm \delta A$$

- For example, the specified magnitude of a resistor is 100 Q with, r a limiting error of ±10Ω. The magnitude of the resistor will be between the limits :

$$A = 100 \pm 10\Omega$$

$$A \geq 90\ \Omega \text{ and } A \leq 110\ \Omega$$

- In other words the manufacturer guarantees that the value of resistance of the resistor lies between 90Ω and 110Ω.

Relative Limiting Error (ε_r)

- The **relative (fractional) error** is *defined as the ratio of the error to the specified magnitude of a quantity*. Therefore,

Relative limiting error

$$\varepsilon_r = \delta A / A1$$

Then rewriting above equation as follows :

$$A = A1 \pm \delta A$$

$$= A1 \pm \varepsilon_r A1$$

$$= A1 (1 \pm \varepsilon_r)$$

Percentage limiting error ($\%\varepsilon_r$)

$$\%\varepsilon_r = \varepsilon_r \times 100$$

Explain the relative limiting error and its effect.

- Thus in selecting instruments, particular care should be taken as regards the range. The values to be measured should, not lie in the lower third of the range. This is particularly important if the accuracy is specified terms of full scale deflection f.s.d. as considerable error (as a percentage of actual value) may occur as is observed from the example 7.2.

- **Example 7.2.** A flow meter working on thermal principles has a guaranteed accuracy of ±5% of full scale reading of 5×10^{-6} m^3/s. The flow measured by this meter is 2.5×10^{-6} m^3/s. Calculate the limiting error in percent. Comment on the results.

- **Example 7.3.** A pressure gauge having a range 100O kN/m^2 has an error of ±1% of full scale deflection. If the true pressure is 100 kN/m^2, what would be the range of readings. Suppose the error is specified as percentage of true value, what would, be the range of the readings.

Combination of Quantities with Limiting Errors

Explain from extra sheet

Propagation of Uncertainties (ωA)

- The uncertainty analysis in measurements when many variates are involved is done on the same basis as is done for error analysis when the results are expressed as standard deviations or probable errors.

Suppose X is a function of several variables,

X = f(x1 , x2 , x3 , ……., xn)

where x1, x2, x3,...xn are independent variables with the same degree of odds.

Let ω_x be the resultant uncertainty and ω_{x1} , ω_{x2} , ω_{x3} ……… ω_{xn} be the uncertainties in the independent variables x1, x2, x3 ………… xn respectively. The uncertainty in the result is given by :

$$\omega X = [((\partial X/\partial x1)^{*}\ \omega_{x1})^2 + \ldots\ldots ((\partial X/\partial xn)^{*}\ \omega_{xn})^2]^{0.5}$$