UNIT 1 CHARACTERISTICS OF NORMAL DISTRIBUTION

Structure

1.0 INTRODUCTION

So far you have learnt in descriptive statistics, how to organise a distribution of scores and how to describe its shape, central value and variation. You have used histogram and frequency polygon to illustrate the shape of a frequency distribution, measures of central tendency to describe the central value and measures of variability to indicate its variation. All these descriptions have gone a long way in providing information about a set of scores, but we also need procedures to describe individual scores or cutting point scores to categorize the entire group of individuals on the basis of their ability or the nature of test paper, which a psychometerician or teacher has used to assess the outcomes of the individual on a certain ability test. For example, suppose a teacher has administered a test designed to appraise the level of achievement and a student has got some score on the test. What did that score mean? The obtained score has some meaning only with respect to other scores either the teacher may be interested to know how many students lie within the certain range of scores? Or how many students are above and below certain referenced score? Or how many students may be assign A, B, C, D etc. grades according to their ability?

To have an answer to such problems, the curve of Bell shape, which is known as Normal curve, and the related distribution of scores, through which the bell shaped curve is obtained, generally known as Normal Distribution, is much helpful.

Thus the present unit presents the concept, characteristics and use of Normal Distributions and Normal Curve, by suitable illustrations and explanations.

1.1 OBJECTIVES

After reading this unit, you will be able to:

- Explain the concept of normal distribution and normal probability curve;
- Draw the normal probability curve on the basis of given normal distribution;
- Explain the theoretical basis of the normal probability curve;
- Elucidate the Characteristics of the normal probability curve and normal distribution;
- Analyse the normal curve obtained on the basis of large number of observations;
- Describe the importance of normal distribution curve in mental and educational measurements;
- Explain the applications of normal curve in mental measurement and educational evaluation;
- Read the table of area under normal probability curve;
- Compare the Non-Normal with normal Distribution and express the causes of divergence from normalcy; and
- Explain the significance of skewness and kurtosis in the mental measurement and educational evaluation.

1.2 NORMAL DISTRIBUTION/ NORMAL PROBABILITY CURVE

1.2.1 Concept of Normal Distribution

Carefully look at the following hypothetical frequency distribution, which a teacher has obtained after examining 150 students of class IX on a Mathematics achievement test.

Table 1.2.1: Frequency distribution of the Mathematics achievement test scores

Class Intervals	Tallies	Frequency
85 – 89	I	1
80 – 84	II	2
75 – 79	IIII	4
70 – 74	𝍸 II	7
65 – 69	𝍸 𝍸	10
60 – 64	𝍸 𝍸 𝍸 I	16
55 – 59	𝍸 𝍸 𝍸 𝍸	20
50 – 54	𝍸 𝍸 𝍸 𝍸 𝍸 𝍸	30
45 – 49	𝍸 𝍸 𝍸 𝍸	20
40 – 44	𝍸 𝍸 𝍸 I	16
35 – 39	𝍸 𝍸	10
30 – 34	𝍸 II	7
25 – 29	IIII	4
20 – 24	II	2
15 – 19	I	1
	Total	**150**

Are you able to find some special trend in the frequencies shown in the column 3 of the above table? Probably yes! The concentration of maximum frequencies (f = 30) lies near a central value of distribution and frequencies gradually tapper off symmetrically on both the sides of this value.

1.2.2 Concept of Normal Curve

Now, suppose if we draw a frequency polygone with the help of above distribution, we will have a curve as shown in the fig. 1.2.1

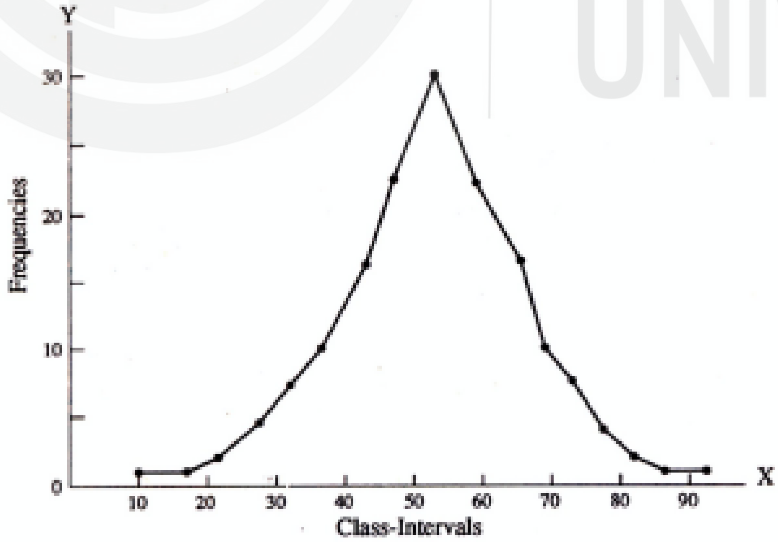

Fig. 1.2.1: Frequency Polygon of the data given in Table 1.2.1

The shape of the curve in Fig. 1.2.1 is just like a 'Bell' and is symmetrical on both the sides.

If you compute the values of Mean, Median and Mode, you will find that these three are approximately the same (M = 52; Md = 52 and Mo = 52)

This Bell shaped curve technically known as Normal Probability Curve or simply Normal Curve and the corresponding frequency distribution of scores, having just the same values of all three measures of central tendency (Mean, Median and Mode) is known as Normal Distribution.

Many variables in the physical (e.g. height, weight, temperature etc.) biological (e.g. age, longevity, blood sugar level and behavioural (e.g. Intelligence; Achievement; Adjustment; Anxiety; Socio-Economic-Status etc.) sciences are normally distributed in the nature. This normal curve has a great significance in mental measurement. Hence to measure such behavioural aspects, the Normal Probability Curve in simple terms Normal Curve worked as reference curve and the unit of measurement is described as σ (Sigma).

1.2.3 Theoretical Base of the Normal Probability Curve

The normal probability curve is based upon the law of Probability (the various games of chance) discovered by French Mathematician Abraham Demoiver (1667-1754). In the eighteenth century, he developed its mathematical equation and graphical representation also.

The law of probability and the normal curve that illust-rates it are based upon the law of chance or the probable occurrence of certain events. When any body of observations conforms to this mathematical form, it can be represented by a bell shaped curve with definite characteristics.

1.2.4 Characteristics or Properties of Normal Probability Curve (NPC)

The characteristics of the normal probability curve are:

1) **The Normal Curve is Symmetrical:** The normal probability curve is symmetrical around it's vertical axis called ordinate. The symmetry about the ordinate at the central point of the curve implies that the size, shape and slope of the curve on one side of the curve is identical to that of the other. In other words the left and right halves to the middle central point are mirror images, as shown in the figure given here.

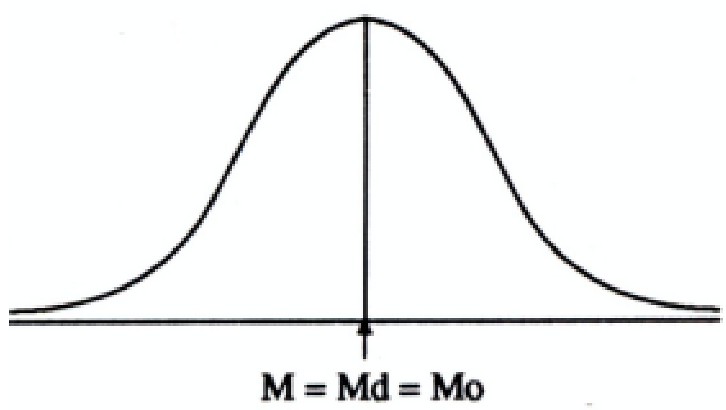

Fig. 1.2.2

2) **The Normal Curve is Unimodel:** Since there is only one maximum point in the curve, thus the normal probability curve is unimodel, i.e. it has only one mode.

3) **The Maximum Ordinate occurs at the Center:** The maximum height of the ordinate always occur at the central point of the curve, that is the mid-point. In the unit normal curve it is equal to 0.3989.

4) **The Normal Curve is Asymptotic to the X Axis:** The normal probability curve approaches the horizontal axis asymptotically; i.e. the curve continues to decrease in height on both ends away from the middle point (the maximum ordinate point); but it never touches the horizontal axis. Therefore its ends extend from minus infinity (- ∞) to plus infinity (+ ∞).

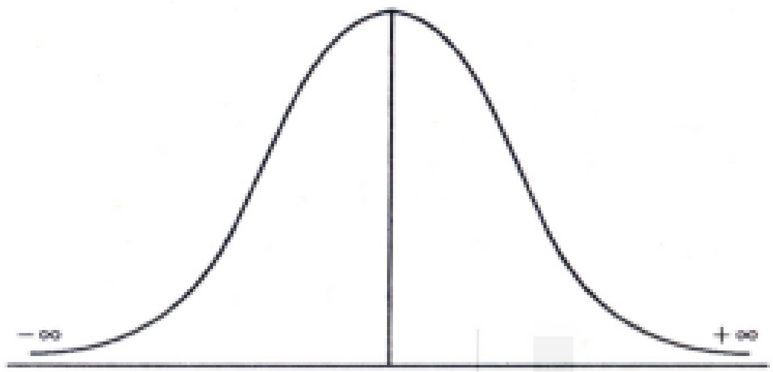

Fig. 1.2.3

5) **The Height of the Curve declines Symmetrically:** In the normal probability curve the height declines symmetrically in either direction from the maximum point.

6) **The Points of Influx occur at point ±1 Standard Deviation (± 1 σ):** The normal curve changes its direction from convex to concave at a point recognised as point of influx. If we draw the perpendiculars from these two points of influx of the curve to the horizontal X axis; touch at a distance one standard deviation unit from above and below the mean (the central point).

7) **The Total Percentage of Area of the Normal Curve within Two Points of Influxation is fixed:** Approximately 68.26% area of the curve lies within the limits of ± 1 standard deviation (± 1 σ) unit from the mean.

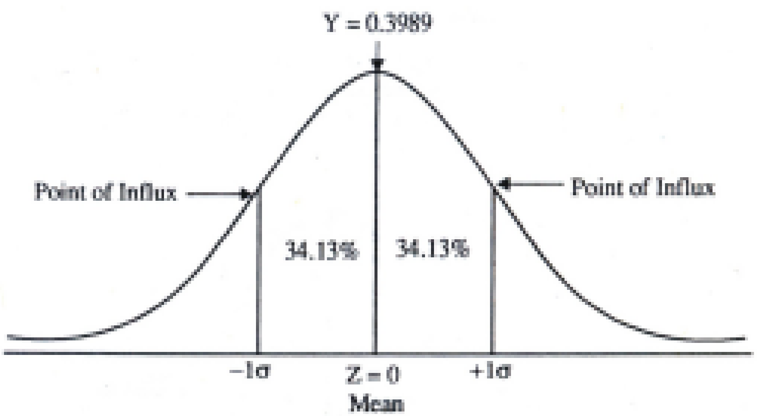

Fig. 1.2.4

8) **The Total Area under Normal Curve may be also considered 100 Percent Probability:** The total area under the normal curve may be considered to approach 100 percent probability; interpreted in terms of standard deviations. The specified area under each unit of standard deviation are shown in this figure.

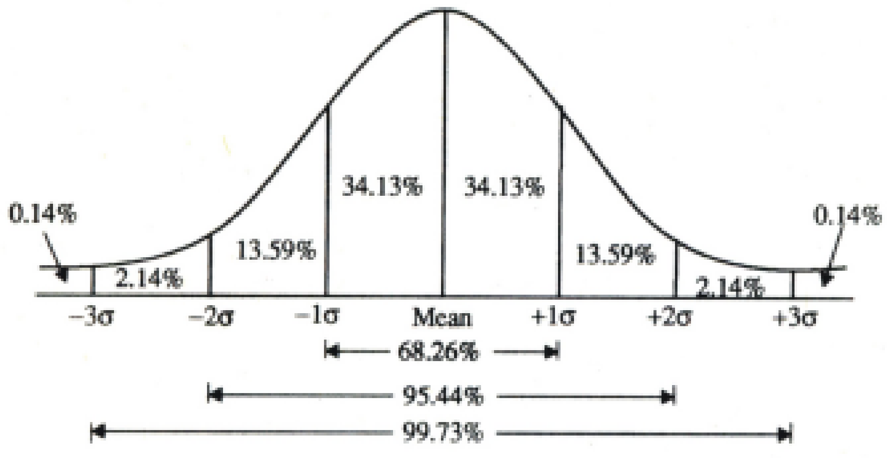

Fig. 1.2.5: The Percentage of the Cases Failing Between Successive Standard Deviation in Normal Distribution

9) **The Normal Curve is Bilateral:** The 50% area of the curve lies to the left side of the maximum central ordinate and 50% of the area lies to the right side. Hence the curve is bilateral.

10) **The Normal Curve is a mathematical model in behavioural Sciences Specially in Mental Measurement:** This curve is used as a measurement scale. The measurement unit of this scale is $\pm$ 1σ (the unit standard deviation).

Self Assessment Questions

1) Define a Normal Probability Curve.

....

....

....

....

2) Write the properties of Normal Distribution.

....

....

....

....

3) Mention the conditions under which the frequency distribution can be approximated to the normal distribution.

....

....

....

....

4) In a distribution what percentage of frequencies are lie in between

(a) -1 σ to + 1 σ

(b) -2 σ to + 2 σ

(c) -3 σ to + 3 σ

....

....

....

....

5) Practically, why are the two ends of normal curve considered closed at the points $\pm 3\ \sigma$ of the base.

....

....

....

....

1.3 INTERPRETATION OF NORMAL CURVE/ NORMAL DISTRIBUTION

What do normal curve/ normal distribution indicate? Normal curve has great significance in the mental measurement and educational evaluation. It gives important information about the trait being measured.

If the frequency polygon of observations or measurements of certain trait is a normal curve, it is a indication that

1) The measured trait is normally distributed in the universe.
2) Most of the cases i.e. individuals are average in the measured trait and their percentage in the total population is about 68.26%.
3) Approximately 15.87% (50-34.13%) of cases are high in the trait measured.
4) Similarly 15.87% of cases are low in the trait measured.
5) The test which is used to measure the trait is good.
6) The test which is used to measure the trait has good discrimination power as it differentiates between poor, average and high ability group individuals.
7) The items of the test used are fairly distributed in terms of difficulty level.

1.4 IMPORTANCE OF NORMAL DISTRIBUTION

The Normal distribution is by far the most used distribution in inferential statistics because of the following reasons:

1) Number of evidences are accumulated to show that normal distribution provides a good fit or describe the frequencies of occurrence of many variable facts in biological statistics, e.g. sex ratio in births, in a country over a number of years. The anthropometrical data, e.g. height, weight, etc. The social and economic data e.g. rate of births, marriages and deaths. In psychological measurements

e.g. Intelligence, perception span, reaction time, adjustment, anxiety etc. In errors of observation in physics, chemistry, astronomy and other physical sciences.

2) The normal distribution is of great value in educational evaluation and educational research, when we make use of mental measurement. It may be noted that normal distribution is not an actual distribution of scores on any test of ability or academic achievement, but is, instead, a mathematical model. The distributions of test scores approach the theoretical normal distribution as a limit, but the fit is rarely ideal and perfect.

1.5 APPLICATIONS/USES OF NORMAL DISTRIBUTION CURVE

There are number of applications of normal curve in the field of psychology as well as educational measurement and evaluation. These are:

i) To determine the percentage of cases (in a normal distribution) within given limits or scores.

ii) To determine the percentage of cases that are above or below a given score or reference point.

iii) To determine the limits of scores which include a given percentage of cases to determine the percentile rank of an individual or a student in his own group.

v) To find out the percentile value of an individual on the basis of his percentile rank.

vi) Dividing a group into sub-groups according to certain ability and assigning the grades.

vii) To compare the two distributions in terms of overlapping.

viii) To determine the relative difficulty of test items.

1.6 TABLE OF AREAS UNDER THE NORMAL PROBABILITY CURVE

How do we use all the above applications of normal curve in mental as well as in educational measurement and evaluation? It is essential first to know about the Table of areas under the normal curve.

The Table 1.6.1 gives the fractional parts of the total area under the normal curve found between the mean and ordinates erected at various σ (sigma) distances from the mean.

The normal probability curve table is generally limited to the areas under unit normal curve with N = 1, σ = 1. In case, when the values of N and σ are different from these, the measurements or scores should be converted into sigma scores (also referred to as standard scores or z scores). The process is as follows :

$$z = \frac{X - M}{\sigma} \quad \text{or} \quad z = \frac{x}{\sigma}$$

In which: z = Standard Score X = Raw Score

M = Mean of X Scores σ = Standard Deviation of X Scores

The table of areas of normal probability curve are then referred to find out the proportion of area between the mean and the z value.

Though the total area under the N.P.C. is 1, but for convenience, the total area under the curve is taken to be 10,000 because of the greater ease with which fractional parts of the total area, may be then calculated.

The first column of the table, x/σ gives distance in tenths of σ measured off on the base line for the normal curve from the mean as origin. In the row, the x/σ distance are given to the second place of the decimal.

To find the number of cases in the normal distribution between the mean, and the ordinate erected at a distance of lσ unit from the mean, we go down the x/σ column until 1.0 is reached and in the next column under .00 we take the entry opposite 1.0, namely 3413. This figure means that 3413 cases in 10,000; or 34.13 percent of the entire area of the curve lies between the mean and lσ. Similarly, if we have to find the percentage of the distribution between the mean and 1 .56 σ, say, we go down the x/σ column to 1.5, then across horizontally to the column headed by .06, and note the entry 44.06. This is the percentage of the total area that lies between the mean and 1.56σ.

Table 1.6.1: Fractional parts of the total area (taken as 10,000) under the normal probability curve, corresponding to distance on the baseline between the mean and successive points laid off from the mean in units of standard deviation.

x/σ	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0000	0040	0080	0120	0160	0199	0239	0279	0319	0359
0.1	0398	0438	0478	0517	0557	0596	0636	0675	0714	0753
0.2	0793	0832	0871	0910	0948	0987	1026	1064	1103	1141
0.3	1179	1217	1255	1293	1331	1368	1406	1443	1480	1517
0.4	1554	1591	1628	1664	1700	1736	1772	1808	1844	1879
0.5	1915	1950	1985	2019	2054	2088	2123	2157	2190	2224
0.6	2257	2291	2324	2457	2389	2422	2454	2486	2517	2549
0.7	2580	2611	2642	2673	2704	2734	2764	2794	2823	2852
0.8	2881	2910	2939	2967	2995	3023	3051	3078	3106	3133
0.9	3159	3186	3212	3238	3264	3290	3315	3340	3365	3389
1.0	3413	3438	3461	3485	3508	3531	3554	3577	3599	3621
1.1	3643	3665	3686	3708	3729	3749	3770	3790	3810	3830
1.2	3849	3869	3889	3907	3925	3944	3962	3980	3997	4015
1.3	4032	4049	4066	4082	4099	4115	4131	4147	4162	4177
1.4	4192	4207	4222	4236	4251	4265	4279	4292	4306	4319
1.5	4332	4345	4357	4370	4383	4394	4406	4418	4429	4441
1.6	4452	4463	4474	4484	4495	4505	4515	4525	4535	4545
1.7	4554	4564	4573	4582	4591	4599	4608	4616	4625	4633
1.8	4641	4649	4656	4664	4671	4678	4686	4693	4699	4706
1.9	4713	4719	4726	4732	4738	4744	4750	4756	4761	4767
2.0	4772	4778	4783	4788	4793	4798	4803	4808	4812	4817
2.1	4821	4826	4830	4834	4838	4842	4846	4850	4854	4857
2.2	4861	4864	4868	4871	4875	4878	4881	4884	4887	4890
2.3	4893	4896	4898	4901	4904	4906	4909	4911	4913	4916
2.4	4918	4920	4922	4925	4927	4929	4931	4932	4934	4936
2.5	4938	4940	4941	4943	4945	4946	4948	4949	4951	4952
2.6	4953	4955	4956	4957	4959	4960	4961	4962	4963	4964
2.7	4965	4966	4967	4968	4969	4970	4971	4972	4973	4974
2.8	4974	4975	4976	4977	4977	4978	4979	4979	4980	4981
2.9	4981	4982	4982	4988	4984	4984	4985	4985	4986	4986
3.0	4986.5	4986.9	4987.4	4987.8	4988.2	4988.6	4988.9	4989.3	4989.7	4990.0
3.1	4990.3	4990.6	4991.0	4991.3	4991.6	4991.8	4992.1	4992.4	4992.6	4992.9
3.2	4993.129									

3.3	4995.166									
3.4	4996.631									
3.5	4997.674									
3.6	4998.409									
3.7	4998.922									
3.8	4999.277									
3.9	4999.519									
4.0	4999.683									
4.5	4999.966									
5.0	4999.997133									

Example: Between the mean and a point 1.38 σ $\left(\frac{x}{\sigma}=1.38\right)$ are found 41.62% of the entire area under the curve.

We have so far considered only σ distances measured in the positive direction from the mean. For this we have taken into account only the right half of the normal curve. Since the curve is symmetrical about the mean, the entries in Table apply to distances measured in the negative direction (to the left) as well as to those measured in the positive direction. If we have to find the percentage of the distribution between mean and -1.28σ, for instance, we take entry 3997 in the column .08, opposite 1.2 in the x/σ column. This entry means that 39.97 percent of the cases in the normal distribution fall between the mean and -1.28σ.

For practical purposes we take the curve to end at points -3σ and +3σ distant from the mean as the normal curve does not actually meet the base line. Table of area under normal probability curve shows that 4986.5 cases lie between mean and ordinate at +3σ. Thus 99.73 percent of the entire distribution, would lie within the limits -3σ and +3σ. The rest 0.27 percent of the distribution beyond ±3σ is considered too small or negligible except where N is very large.

1.7 POINTS TO BE KEPT IN MIND WHILE CONSULTING TABLE OF AREA UNDER NORMAL PROBABILITY CURVE

The following points are to be kept in mind to avoid errors, while consulting the N.P.C. Table.

1) Every given score or observation must be converted into standard measure i.e. Z score, by using the following formula:

   $$z = \frac{X - M}{\sigma}$$

2) The mean of the curve is always the reference point, and all the values of areas are given in terms of distances from mean which is zero.

3) The area in terms of proportion can be converted into percentage, and

4) While consulting the table, absolute values of z should be taken. However, a negative value of z shows that the scores and the area lie below the mean and this fact should be kept in mind while doing further calculation on the area. A positive value of z shows that the score lies above the mean *i.e.* right side.

Self Assessment Questions

i) What formula is to use to convert raw score X into standard score i.e. z score.

....

....

ii) What is the reference point on the normal probability curve.

iii) Mean value of the z scores is ____________

iv) The value of standard deviation of z scores is ___________

v) The total area under the N.P.C. is always ___________

vi) The negative value of z scores shows that ___________

vii) The positive value of z scores shows that _________

1.8 PRACTICAL PROBLEMS RELATED TO APPLICATION OF THE NORMAL PROBABILITY CURVE

Under the caption 1.5 you have studied the Application of normal Distribution/ Normal Curve in mental and educational measurements. Now how the practical problems related to these application are solved you go through the following examples carefully and thoughroly.

1) **To determine the percentage of cases in a normal distribution within given limits of scores.**

Often a psychometrician or psychology teacher is interested to know the number of cases or individuals that lie in between two points or two limits. For example, a teacher may be interested as to how many students of his class got marks in between 60% and 70% in the annual examination, or he may be interested in how many students of his got marks above 80%.

Example 1

An adjustment test was administered on a sample of 500 students of class VIII. The mean of the adjustment scores of the total sample obtained was 40 and standard deviation obtained was 8, what percentage of cases lie between the score 36 and 48, if the distribution of adjustment scores is normal in the universe.

Solution:

In the problem it is given that

N = 500

M = 40

$\sigma = 8$

We have to find out the total % of the students who obtained score in between 36 and 48 on the adjustment test.

To find the required percentage of cases, first we have to find out the z scores for the raw scores (X) 36 and 48, by using the formula.

$$z = \frac{X - M}{\sigma}$$

$\therefore$ z score for raw score 36 is

$$z_1 = \frac{36-40}{8} =$$

or $\quad z_1 = -0.5\ \sigma$

Similarly z score for raw score 40 is

$$z_2 = \frac{48-40}{8} =$$

or $\quad z_2 = +1\ \sigma$

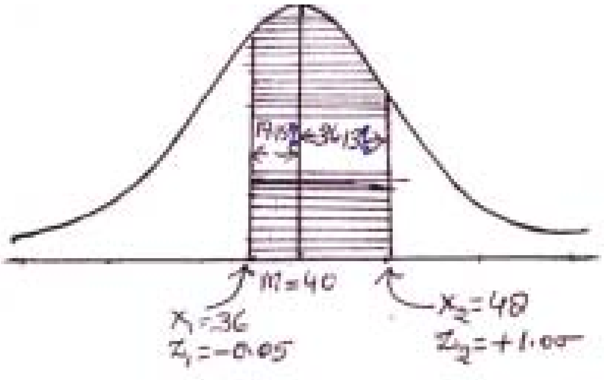

Fig. 1.8.1

According to table of area under Normal Probability curve (N.P.C.) *i.e.* Table No. 1.6.1 the total area of the curve lie in between M to +1σ is 34.13 and in between M to -0.56 is 19.15.

$\therefore$ The total area of the curve in between -0.5 σ to +1 σ is 19.15 + 34.13 = 53.28

Thus the total percentage of students who got scores in between 36 and 48 on the adjustment test is 53.28 (Ans.)

Example 2

A reading ability test was administered on the sample of 200 cases studying in IX class. The mean and standard deviation of the reading ability test score was obtained 60 and 10 respectively. Find how many cases lie in between the scores 40 and 70. Assume that reading ability scores are normally distributed.

Solution:

Given $\quad N = 200$

$M = 60$

$\sigma = 10$

$X_1 = 40$ and

$X_2 = 70$

To find out: The total no. of cases in between the two scores 40 and 70.

To find the required no. of cases, first we have to find out the total percentage of cases lie in between Mean and 40 and mean and 70. See the Fig. 1.8.2 For the purpose, first the given raw scores (40 & 70) should be converted into z scores by using the formula

$$z = \frac{X-M}{\sigma}$$

$$\therefore \quad z_1 = \frac{40-60}{10} =$$

or $\quad z_1 = -\ 2\sigma$

Similarly $z_2 = \frac{70-60}{10} =$

or $z_2 = +1\sigma$

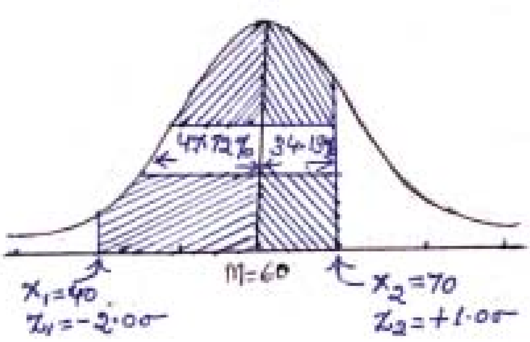

Fig. 1.8.2

According to Table 1.6.1 the area of the curve in between M and -2σ is 47.72% and in between M and +1σ is 34.13%.

∴ The total area of the curve in between -2σ to +1σ is = 47.72 + 34.13 = 81.85%

Therefore, the total no. of cases in between the two scores 40 and 70 are =

$$\frac{81.85 \times 200}{100} = 163.7 \text{ or } 164$$

Thus total no. of cases who got scores in between 40 and 70 are = 164. (Ans.)

2) **To determine the percentage of cases lie above or below a given score or reference point.**

Example 3

An intelligence test was administered on a group of 500 cases of class V. The mean I.Q. of the students was found 100 and the S.D. of the I.Q. scores was 16. Find how many students of class V having the I.Q. below 80 and above 120.

Solution:

Given M = 100, σ = 16, X_1 = 80 and X_2 = 120

To find out : (i) The total no. of cases below 80

(ii) The total no. of cases above 120

To find the required no. of cases first we have to find z scores of the raw scores X_1 = 80 and X_2 = 120 by using the formula

$$z = \frac{X-M}{\sigma}$$

$$z_1 = \frac{80-100}{16} = -\frac{20}{16}$$

or $z_1 = -1.25\ \sigma$

Similarly,

$$z_2 = \frac{120-100}{16} = +\frac{20}{16}$$

or $z_2 = +1.25\ \sigma$

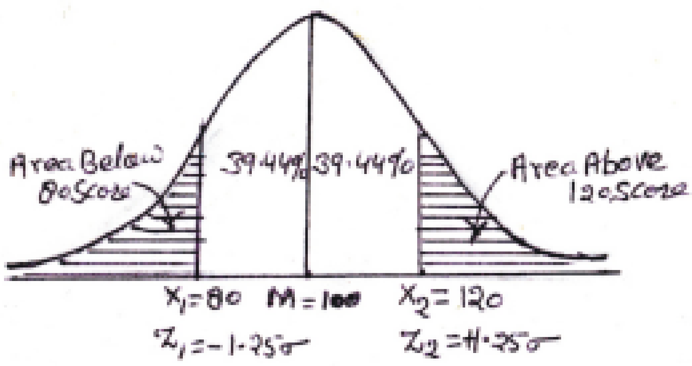

Fig. 1.8.3

According to NPC table (Table 1.6.1) the total percentage of area of the curve lie in between Mean to 1.25 σ is = 39.44.

According to the properties of N.P.C. the 50% area lies below to the mean *i.e.* in left side and 50% area lie above to the mean i.e. in right side.

Thus the total area of NPC curve below M = (100) is = 50 – 39.44 = 10.56

Similarly the total area of NPC curve above M = (100) is = 50 – 39.44 = 10.56

Therefore total cases below to the I.Q. 80 = = 52.8 = 53 Appox.

Similarly Total cases above to the I.Q. 120 = = 52.8 = 53 Appox.

Thus in the group of 500 students of V class there are total 53 students having I.Q. below 80. Similarly there are 53 students who have I.Q. above 120. (Ans.)

3) **To determine the limits of scores which includes a given percentage of cases**

Some time a psychometerician or a teacher is interested to know the limits of the scores in which a specified group of individuals lies. To understand, read the following example-4 and its solution.

Example 4

An achievement test of mathematics was administered on a group of 75 students of class VIII. The value of mean and standard deviation was found 50 and 10 respectively. Find limits of the scores in which middle 60% students lies.

Solution:

Given that, N = 75, M = 50, σ = 10

To find out: Value of the limits of middle 60% cases i.e. X_1 and X_2

As per given condition (middle 60% cases), 30%-30% cases lie left and right to the mean value of the group (see the fig. 1.8.4.)

According to the formula

$$z = \frac{X - M}{\sigma}$$

If the value of M, σ and z is known, the value of X can be find out. In the given problem the value of M and σ are given. We can find out the value of z with the help of the NPC Table No. 1.6.1 as the area of the curve situated right and left to the mean (30%-30% respectively) is also given.

According to the table (1.6.1) the value of z_1 and z_2 of the 30% area is $\pm$ 0.84σ

Therefore by using formula

$$z_1 = \frac{X_1 - M}{\sigma}$$

$$-0.84 = \frac{X_1 - 50}{10}$$

or $X_1 = 50 - 0.84 \times 10$

$= 41.60$ or 42

Similarly,

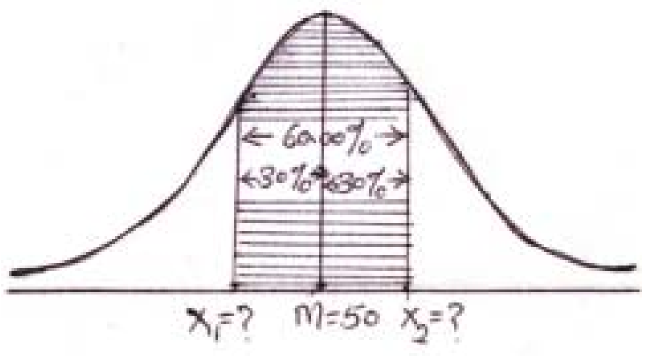

Fig. 1.8.4

$$z_2 = \frac{X_2 - M}{\sigma}$$

$$-0.84 = \frac{X_2 - 50}{10}$$

or $X_2 = 50 + 0.84 \times 10$

$= 58.4$ or 58

Thus $X_1 = 42$

$X_2 = 58$

Therefore, the middle 60% cases of the entire group (75, students) got marks on achievement test of mathematics in between 42 – 58. Ans.

Self Assessment Questions

The observation given in the example 4, i.e. M = 50 and S.D. (σ) = 10

1) Find the limits of the scores middle 30% cases

....

....

2) Find the limits of the scores middle 75% cases

....

....

3) Find the limits of the scores middle 50% cases

....

....

4) **To determine the percentage rank of the individual in his group.**

The percentile rank is defined as the percentage of cases lie below to a certain score (X) or a point.

Some time a psychologist or a teacher is interested to know the position of an individual or a student in his own group on the bases of the trait is measured (for more clarification go through the following example carefully)

Example 5

In a group of 60 students of class X, Sumit got 75% marks in board examination. If the mean of whole class marks is 50 and S.D. is 10. Find the percentile rank of the Sumit in the class.

Solution:

See the fig. 1.8.5. and pay the attention to the definition of percentile given above carefully.

It is clear from the fig. that we have to find out the total percentage of cases (i.e. the area of N.P.C.) lie below to the point X = 75 (See Fig. 1.8.5.)

To find the total required area (shaded part) of the curve, it is essential first to know the area of the curve lie in between the points 50 and 75.

This area can be determined very easily, by taking up the help of N.P.C. Table, i.e. Table No. 1.6.1., if we know the value of z of score 75.

According to the formula

$$z = \frac{X-M}{\sigma}$$

$$z = \frac{75-50}{10} = \frac{25}{10}$$

or $= +2.50\ \sigma$

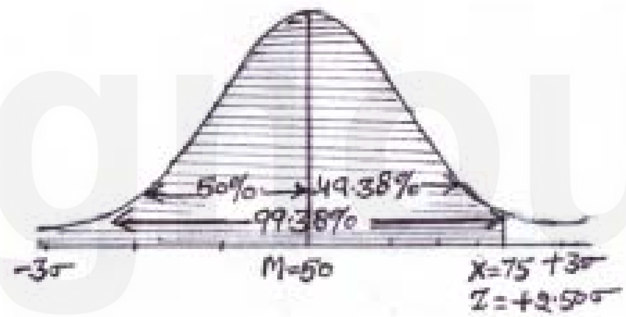

Fig. 1.8.5

According to NPC Table (Table No. 1.6.1) the area of the curve lies M and +2.50 σ is 49.387.

In the present problem we have determined 49.38% area lies right to the mean and 50% area lies to the left of the Mean. (According to the properties of NPC see caption 1.2.4 property no. 9)

Thus according to the definition of percentile the total area of the curve lies below to the point X = 75 is

= 50 + 49.38%

= 99.38% or 99% Approx.

Therefore the percentile rank of the Sumit in the class is 99th. In other words Sumit is the topper student in the class, remaining 99% students lie below to him. (Ans.)

Self Assessment Questions

In a test of 200 items, each correct item has 1 mark.

If M = 100, σ = 10

1) Find the position of Rohit in the group who secured 85 marks on the test.

....

....

2) Find the percentile rank of Sunita she got 130 marks on the test.

....

....

5) **To find out the percentile value of an individual's percentile rank.**

Some time we are interested to know that the person or an individual having a specific percentile rank in the group, than what is the percentage of score he got on the test paper. To understand, go through the following example and its solution –

Example 6

An intelligence test was administered on a large group of student of class VIII. The mean and standard deviation of the scores was obtained 65 and 15 respectively. On the basis intelligence test if the Ramesh's percentile rank in the class is 80, find what is the score of the Ramesh, he got on the test?

Solution:

Given : M = 65, σ = 15, and PR = 80

To find out : The value of P_{80}

Look at the Fig. No. 1.8.6., as per definition of percentile rank, the 30% area of the curve lie from mean to the point P_{80} and 50% are lie to the left side of the mean.

The z value of the 30% area of the curve lie in between M and P_{80} is = +0.85 σ

(Table No. 1.16)

We know that $z = \dfrac{X - M}{\sigma}$

or $+\,0.85 = \dfrac{X - 65}{15}$

or $X = 65 + 15 \times 0.85$

$= 65 + 12.75$

$= 77.75$ or 78 Approx.

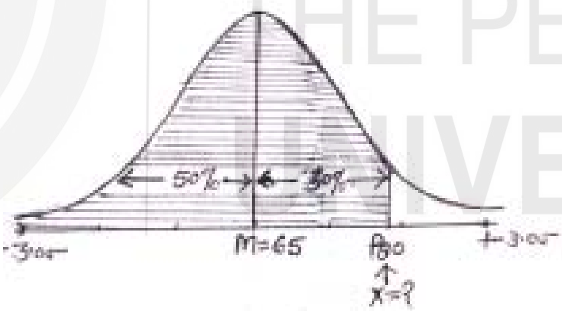

Fig. 1.8.6

Thus Ramesh's intelligence score on the test is = 78 (Ans.)

Self Assessment Questions

1) If M = 100, σ = 10

Find the values of

i) P_{75} = ______________________

ii) P_{10} = ______________________

iii) P_{50} = ______________________

iv) P_{80} = ______________________

6) **Dividing a group of individuals into sub-group according to the level of ability or a certain trait. If the trait or ability is normally distributed in the universe.**

Some time we are making qualitative evaluation of the person or an individual on the basis of trait or ability, and assign them grades like A, B, C, D, E etc. or 1st grade 2nd grade, 3rd grade etc. or High, Average or Low. For example a company evaluate their salesman as A grade, B grade and C grade salesman. A teacher provides A, B, C etc. grades to his students on the basis of their performance in the examination. A psychologist may classify a group of person on the basis of their adjustment as highly adjusted, Average and poorly adjusted. In such conditions, always there is a question that how many persons or individuals, we have to provide A, B, C, D and E etc. grades to the individuals and categorize them in different groups.

For further clarification go through the following examples:

Example 7

A company wants to classify the group of salesman into four categories as Excellent, Good, Average and Poor on the basis of the sale of a product of the company, to provide incentive to them. If the number of salesman in the company is 100, their average sale of the product per week is 10,00,000 Rs. and standard deviation is Rs. 500/-. Find the number of salesman to place as Excellent, Good, Average and Poor.

Solution:

As per property of the N.P.C. we know that total area of the curve is 6σ over a range of -3σ to +3σ.

According to the problem, the total area of the curve is divided into four categories.

Therefore area of each category is 6σ/4 = ± 1.5σ. It means the distance of each category from the mean on the curve is 1.5σ respectively.

The distance of each category is shown in the figure 1.8.7

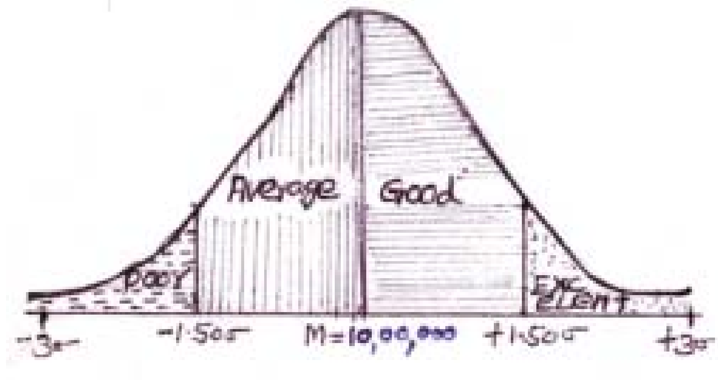

Fig. 1.8.7

i) Total % of salesman in "Good" category

According to N.P.C. Table (Table No. 1.6.1), the

Total area of the curve lies in between M an +1.5σ is = 43.32%

∴ The total % of salesman in "Good" category is 43.32%

ii) Total % of salesman in "Average" category

Total area of the curve lies in between Mean and -1.5σ is also = 43.32%

$\therefore$ The total % of salesman in Average category is = 43.32

iii) Total % of salesman in "Excellent" category

The total area of the curve from M to $+3\sigma$ and above is

= 50% .... (As per properties of Normal Curve)

$\therefore$ The total % of salesman in the category Excellent is = 50 – 43.32 = 6.68%

iv) Total % of salesman in "Poor" category

The total area of the curve from M to -3σ and below is

= 50% .... (As per properties of Normal Curve)

$\therefore$ The total % of the salesman in the poor category is = 50 – 43.32 = 6.68%

Thus,

i) The number of salesman should place in "Excellent" category

$$= \frac{6.68 \times 100}{100} = 6.68 \text{ or } 7$$

ii) The number of salesman should place in "Good" category

$$= \frac{43.32 \times 100}{100} = 43.32 \text{ or } 43$$

iii) The number of salesman should place in "Average" category

$$= \frac{43.32 \times 100}{100} = 43.32 \text{ or } 43$$

iv) The number of salesman should place in "Poor" category

$$= \frac{6.68 \times 100}{100} = 6.68 \text{ or } 7$$

Total = 100 (Ans.)

Self Assessment Questions

In the above example no. 7 if the salesman are categorised into six categories as excellent, v. good, good average, poor and v. poor. Find the number of salesman in each category as per their sales ability.

....

....

Example 8

A group of 1000 applicant's who wishes to take admission in a psychology course. The selection committee decided to classify the entire group into five sub-categories A, B, C, D and E according to their academic ability of last qualifying examination.

If the range of ability being equal in each sub category, calculate the number of applicants that can be placed in groups ABCD and E.

Solution:

Given: N = 1000

To find out: The 1000 cases to be categorised into five categories A, B, C, D, and E.

We know that the base line of a normal distribution curve is considered extend from -3σ to +3σ that is range of 6σ.

Dividing this range by 5 (the five subgroups) to obtain σ distance of each category, i.e. the z value of the cutting point of each category (see the fig. given below)

$$\therefore\ z = \frac{6\sigma}{5} = \pm\ 1.20\ \sigma$$

(It is to be noted here that the entire group of 1000 cases is divided into five categories. The number of subgroups is odd number. In such condition the middle group or middle category (c) will lie equally to the centre i.e. M of the distribution of scores. In other words the number of cases of "c" category or middle category remain half to the left area of the curve from the point of mean and half of the right area of the curve from the mean.

$$\therefore \quad \text{the limits of "c" category is} = \frac{1.2\sigma}{2} = \pm\ 0.60\ \sigma$$

i.e. the "c" category will remain on NPC curve in between the two limits -0.6 σ to +0.6 σ

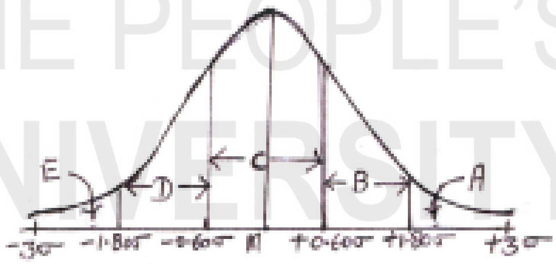

Fig. 1.8.8

Now,

The limits of B category

Lower limit = +0.6 σ

and Upper limit = 0.60 σ + 1.20 σ

or = +1.80 σ

The limits of A category

Lower limit = + 1.8 σ

and Upper limit = + 3 σ and above

Similarly, the limits of D category

Upper limit = - 0.6 σ

Lower limit = (- 0.60 σ) + (-1.20 σ)

or = -1.80 σ

The limits of E category

Upper limit = -1.8 σ

Lower limit = -3σ and below

(For limits of each category see the fig. 1.8.8 carefully)

i) The total % area of the NPC for A category

According to NPC Table (1.6.1) the total % of area in between

Mean to +1.80 σ is = 46.41

∴ The total % of area of the NPC for A category is = 50 – 46.41 = 3.59

ii) The total % Area of the NPC for B category –

According to NPC Table (1.6.1) the total % of Area in between

Mean and + 0.60 σ is = 22.57

∴ The total % area of NPC for B category is = 46.41 – 22.57 = 23.84

iii) The total % area of the NPC for C category –

According to NPC table the total % area of NPC in between

M and + 0.06 σ is = 22.57

Similarly the total % area of NPC in between

M and – 0.06 σ is also = 22.57

∴ The total % area of NPC for C category is = 22.57 + 22.57 = 45.14

iv) In similar way the total % area of NPC for D category is = 23.84

v) The total % area of NPC for E category is = 3.59

Thus the total number of applicants (N = 1000) in –

A category is $= \dfrac{3.59 \times 1000}{100} = 35.9 \quad = 36$

B category is $= \dfrac{23.84 \times 1000}{100} = 238.4 \quad = 238$

C category is $= \dfrac{45.14 \times 1000}{100} = 451.4 \quad = 452$

D category is $= \dfrac{23.84 \times 1000}{100} = 238.4 \quad = 238$

E category is $= \dfrac{3.59 \times 1000}{100} = 35.9 \quad = 36$

Total= 1000 (Ans.)

Self Assessment Questions

1) In the example 8 if the total applicants are categorised into three categories. Find how many applicants will be the categories A, B and C?

....

....

7) **To compare the two distributions in terms of overlapping.**

Example 9

A numerical ability test was administered on 300 graduate boys and 200 graduate girls. The boys Mean score is 26 with S.D. (σ) of 4. The girls' mean. Mean score is 28 with a σ 8. Find the total number of boys who exceed the mean of the girls and total number of girls who got score below to the mean of boys.

Solution:

Given: For Boys, N = 300, M = 26 and σ = 6

For Girls, N = 200, M = 28 and σ = 8

To find: 1- Number of boys who exceed the mean of girls

2- Number of girls who scored below to the mean of boys

As per given conditions, first we have to find the number of cases above the point 28

(The mean of the numerical ability scores of girls) by considering M=26 and σ=6

Second, we to find no. of cases below to the point 26 (The mean score of the boys), by considering M = 28 and 5 – 8 (see the fig. 1.8.9 given below carefully)

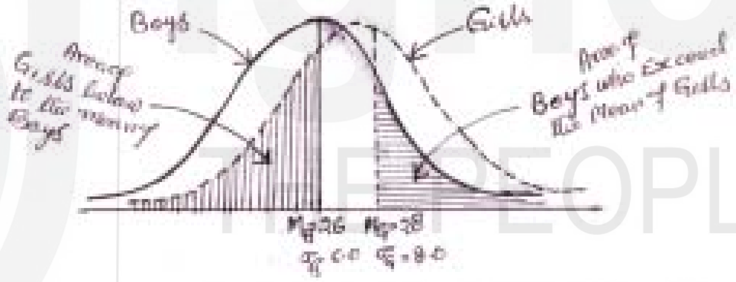

Fig. 1.8.9

1) The z score of X (28) is = $\dfrac{28-26}{6} = \dfrac{2}{6}$

or = + 0.33 σ

According to NPC Table (1.6.1) the total % of area of the NPC from M = 26 to + 0.33 σ is = 12.93

$\therefore$ The total % of cases above to the point 28 is = 50 – 12.93 = 37.07

Thus the total number of boys above to the point 28 (mean of the girls) is

$$= \frac{37.07 \times 300}{100} = 111.21 = 111$$

2) The z score of X = 26 is = $\dfrac{26-28}{8} = \dfrac{2}{8}$ = - 0.25 σ

According to the NPC table the total % of area of the curve in between M = 28 and -0.25 σ is = 9.87

$\therefore$ Total % of cases below to the point 26 is = 50 – 9.87 = 40.13

Thus the total number of girls below to the point 26 (mean of the boys) is

$$= \frac{40.13 \times 200}{100} = 80.26 = 80$$

Therefore,

1) The total number of boys who exceed the mean of the girls in numerical ability is = 111

2) The total number of girls who are below to the mean of the boys is = 88 (Ans.)

Self Assessment Questions
1) In the example given above (Example 9) find.
i) Number of boys between the two means 26 and 28 __________
ii) Number of girls between the two means 26 and 28 __________
iii) Number of boys below to the mean of girls __________
iv) Number of girls above to the mean of boys __________
v) Number of boys above to the Md of girls which is 28.20 __________
vi) Number of girls exceed to the Md of the boys which is 26.20 __________

8) **To determine the relative difficulty of a test items:**

Example 10

In a mathematics achievement test ment for 10th standard class, Q.No. 1, 2 and 3 are solved by the students 60%, 30% and 10% respectively find the relative difficulty level of each Q. Assume that solving capacity of the students is normally distributed in the universe.

Given: The percentage of the students who are solving the test items (Qs) of a question paper correctly.

To Find: The relative difficulty level of each item of the test paper given.

Solution:

First of all we shall mark the relative position of test items on the basis of percentage of students solving the items successfully on the NPC scale.

Q.No.3 of the test paper is correctly solved by the 10% students only. It means 90% students unable to attend the Q.No. 3. On the NPC scale, these 10% cases lies extreme to the right side of the mean (see the fig. given below). Similarly 30% students who are solving Q.No. 2 correctly also lying to the right side of the curve. While the 60% students who are solving Q.No. 1 correctly are lying left side of the N.P.C. curve.

Now, we have to find out the z value of the cut point of the each item (Q.No.) on the NPC base line

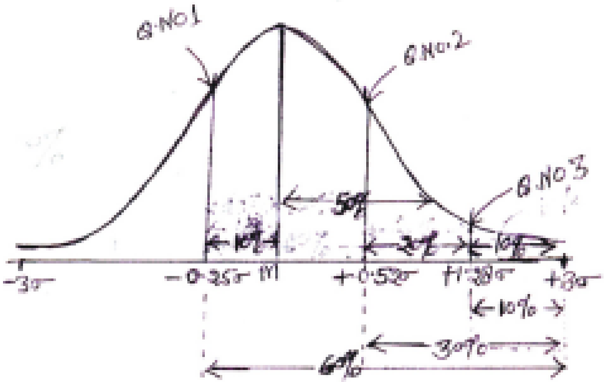

Fig. 1.8.10

i) The z value of the cut point of Q.No. 3

The total percentage of cases lie in between mean and cut point of Q.No. 3 is = (50% - 10%) in right half of NPC

∴ The z value of the right 40% of area of the NPC is = 11.28 σ

ii) The z value of the cut point of Q.No.2

The total percentage of cases lie between the mean and cut point of Q.No. 2 is = 20% (50% - 30%) in right half of NPC

∴ The z value of the right 20% area of the NPC is = + 0.52 σ

iii) The z value of the cut point of Q.No. 3

The total percentage of cases lie between the mean and cut point of Q.No. 3 is = (60% - 50%) in left half of NPC

∴ The z value of the left of 10% of area = - 0.25 σ

Therefore corresponding z value of each item (Q) passed by the students is

Item (Q.No.)	Passed By	z value	Z difference
3	10%	+ 1.28 σ	-
2	30%	+ 0.52 σ	0.76 σ
1	60%	- 0.25 σ	0.77 σ

We may now compare the three questions of the mathematics achievement test, Q.No. 1 has a difficulty value of 0.76 σ higher than the Q.No. 2. Similarly the Q.No. 2 has a difficulty value of 0.77 σ higher than the Q.No. 3. Thus the Q.No. 1, 2 and 3 of the mathematics achievement test are the good items having equal level of difficulty and are quite discriminative. (Ans.)

Self Assessment Question

1) The three test items 1, 2 and 3 of an ability test are solved by 10%, 20% and 30% respectively. What are the relative difficulty values of these items?

....

....

....

1.9 DIVERGENCE IN NORMALITY (THE NON-NORMAL DISTRIBUTION)

In a frequency polygon or histogram of test scores, usually the first thing that strikes one is the symmetry or lack of it in the shape of the curve. In the normal curve model, the mean, the median and the mode all coincide and there is perfect balance between the right and left halves of the curve. Generally two types of divergence occur in the normal curve.

1) Skewness

2) Kurtosis

1) **Skewnes:** A distribution is said to be "skewed" when the mean and median fall at different points in the distribution and the balance *i.e.* the point of center of gravity is shifted to one side or the other to left or right. In a normal distribution the mean equals, the median exactly and the skewness is of course zero ($S_K = 0$).

   There are two types of skewness which appear in the normal curve.

   a) **Negative Skewness :** Distribution said to be skewed negatively or to the left when scores are massed at the high end of the scale, *i.e.* the right side of the curve are spread out more gradually toward the low end *i.e.* the left side of the curve. In negatively skewed distribution the value of median will be higher than that of the value of the mean.

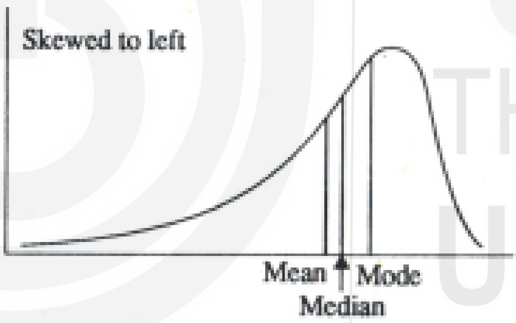

Fig. 1.9.1: Negative Skewness

   b) **Positive Skewness:** Distributions are skewed positively or to the right, when scores are massed at the low; *i.e.* the left end of the scale, and are spread out gradually toward the high or right end as shown in the fig.

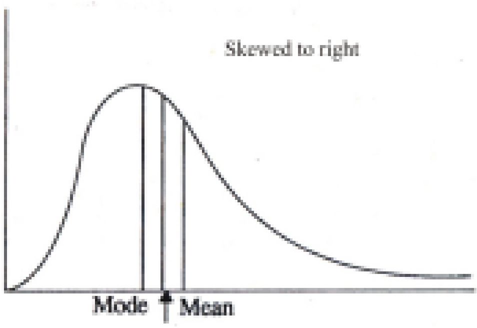

Fig. 1.9.2: Negative Skewness

2) **Kurtosis:** The term kurtosis refers to (the divergence) in the height of the curve, specially in the peakness. There are two types of divergence in the peakness of the curve

a) **Leptokurtosis:** Suppose you have a normal curve which is made up of a steel wire. If you push both the ends of the wire curve together. What would happen in the shape of the curve? Probably your answer may be that by pressing both the ends of the wire curve, the curve become more peeked *i.e.* its top become more narrow than the normal curve and scatterdness in the scores or area of the curve shrink towards the center.

Thus in a Leptokurtic distribution, the frequency distribution curve is more peaked than to the normal distribution curve.

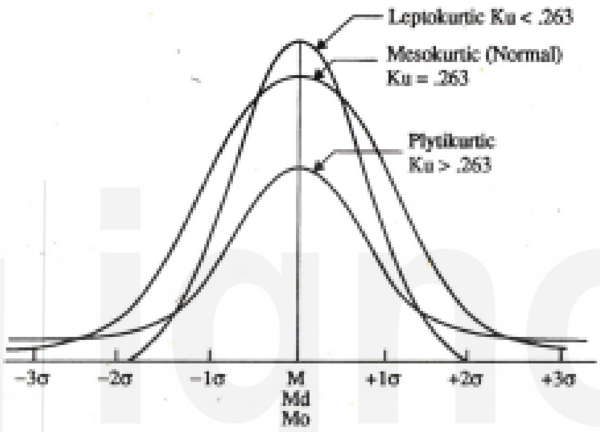

Fig. 1.9.3: Kurtosis in the Normal Curve

b) **Platykurtosis:** Now suppose we put a heavy pressure on the top of the wire made normal curve. What would be the change in the, shape of the curve? Probably you may say that the top of the curve become more flat than to the normal.

Thus a distribution of flatter Peak than to the normal is known Platykurtosis distribution.

When the distribution and related curve is normal, the vain of kurtosis is 0.263 (KU = 0.263). If the value of the KU is greater than 0.263, the distribution and related curve obtained will be platykurtic. When the value of KU is less than 0.263, the distribution and related curve obtained will be Leptokurtic.

1.10 FACTORS CAUSING DIVERGENCE IN THE NORMAL DISTRIBUTION /NORMAL CURVE

The reasons on why distribution exhibit skewness and kurtosis are numerous and often complex, but a careful analysis of the data will often permit the common causes of asymmetry. Some of common causes are –

1) **Selection of the Sample:** Selection of the subjects (individuals) produce skewness and kurtosis in the distribution. If the sample size is small or sample is biased one, skewness is possible in the distribution of scores obtained on the basis of selected sample or group of individuals.

If the scores made by small and homogeneous group are likely to yield narrow and leptokurtic distribution. Scores from small and highly hetrogeneous groups yield platykurtic distribution.

2) **Unsuitable or Poorly Made Tests:** If the measuring tool or test is unappropriate, or poorly made, the asymmetry is possible in the distribution of scores. If a test is too easy, scores will pile up at the high end of the scale, whereas the test is too hard, scores will pile up at the low end of the scale.

3) **The Trait being Measured is Non-Normal:** Skewness or Kurtosis or both will appear when there is a real lack of normality in the trait being measured, e.g. interest, attitude, suggestibility, deaths in a old age or early childhood due to certain degenerative diseases etc.

4) **Errors in the Construction and Administration of Tests:** The unstandardised with poor item-analysis test may cause asymmetry in the distribution of the scores. Similarly, while administrating the test, the unclear instructions – Error in timings, Errors in the scoring, practice and motivation to complete the test all the these factors may cause skewness in the distribution.

Self Assessment Questions

1) Define the following:

a) Skewness

....

....

b) Negative and Positive Skewness

....

....

c) Kurtosis

....

....

d) Platykurtosis

....

....

e) Leptokurtosis

....

....

2) In case of normal distribution what should be the value of skewness.

....

....

....

....

3) In case of normal distribution what should be the value of Kurtosis.

....

....

....

....

4) What is the significance of the knowledge of skewness and kurtosis to a school teacher?

....

....

....

....

1.11 MEASURING DIVERGENCE IN THE NORMAL DISTRIBUTION / NORMAL CURVE

In psychology and education the divergence in normal distribution/normal curve have a significant role in construction of the ability and mental tests and to test the representativeness of a sample taken from a large population. Further the divergence in the distribution of scores or measurements obtained of a certain population reflects some important information about the trait of population measured. Thus there is a need to measure the two divergence i.e. skewness and kurtosis of the distribution of the scores.

1.11.1 Measuring Skewness

There are two methods to study the skewness in a distribution.

i) Observation Method

ii) Statistical Method

i) **Observation Method:** There is a simple method of detecting the directions of skewness by the inspection of frequency polygon prepared on the basis of the scores obtained regarding a trait of the population or a sample drawn from a population.

Looking at the tails of the frequency polygon of the distribution obtained, if longer tail of the curve is towards the higher value or upper side or right side to the centre or mean, the skewness is positive. If the longer tail is towards the lower values or lower side or left to the mean, the skewness is negative.

ii) **Statistical Method:** To know the skewness in the distribution we may also use the statistical method. For the purpose we use measures of central tendency, specifically mean and median values and use the following formula.

$$S_K = \frac{3(Mean - Median)}{\sigma}$$

Another measure of skewness based on percentile values is as under

$$S_K = \frac{(P_{90} - P_{10})}{2} - P_{50}$$

Here, it is to be kept in mind that the above two measures are not mathematically equivalent. A normal curve has the value of $S_K = 0$. Deviations from normality can be negative and positively direction leading to negatively skewed and positively skewed distributions respectively.

1.11.2 Measuring Kurtosis

For judging whether a distribution lacks normal symmetry or peakedness; it may detected by inspection of the frequency polygon obtained. If a peak of curve is thin and sides are narrow to the centre, the distribution is leptokurtic and if the peak of the frequency distribution is too flat and sides of the curve are deviating from the centre towards ±4σ or ±5σ than the distribution is platikurtic.

Kurtosis can be measured by following formula using percentile values.

$$K_U = \frac{Q}{P_{90} - P_{10}}$$

where Q = quartile deviation i.e.

P_{10} = 10th percentile

P_{90} = 90th percentile

A normal distribution has $K_U = 0.263$. If the value of K_U is less than 0.263 ($K_U < 0.263$), the distribution is leptokurtic and if K_U is greater than 0.263 ($K_U > 0.263$), the distribution is platykurtic.

Self Assessment Questions

1) How we can instantly study the skewness in a distribution.

....

....

....

....

2) What is the formula to measure skewness in a distribution?

....

....

....

....

3) What indicates the kurtosis of a distribution?

....

....

....

....

4) What formula is used to calculate the value of kurtosis in a distribution?

....

....

....

....

5) How we decide that a distribution is leptokurtic or platykurtic?

....

....

....

....

1.12 LET US SUM UP

The normal distribution is a very important concept in the behavioural sciences because many variables used in behavioural research are assumed to be normally distributed.

In behavioural science each variable has a specific mean and standard deviation, there is a family of normal distribution rather than just a single distribution. However, if we know the mean and standard deviation for any normal distribution we can transform it into the standard normal distribution. The standard normal distribution is the normal distribution in standard score (z) form with mean equal to 0 and standard deviation equal to 1.

Normal curve is much helpful in psychological and educational measurement and educational evaluation. It provides relative positioning of the individual in a group. It can be also used as a scale of measurement in behavioural sciences.

The normal distribution is a significant tool in the hands of teacher and researcher of psychology and education. Through which he can decide the nature of the distribution of the scores obtained on the basis of measured variable. Also he can decide about his own scoring process which is very lenient or hard; he can Judge the difficulty level of the test items in the question paper and finally he may know about his class, whether it is homogeneous to the ability measured or it is hetrogeneous one.

1.13 UNIT END QUESTIONS

1) Take some frequency distributions and prepare the frequency polygons. Study the normalcy in the distribution. If you will obtained non-normal distribution, determine the type of skewness and kurtosis. Also list down the probable causes associated to the non-normal distribution.

2) Collect the annual examination marks of various subjects of any class and study the nature of distribution of scores of each subject. Also determine the difficulty level of the question papers of each subject.

3) Determine which variables related to cognitive and affective domain of behaviour are normally distributed.

4) As a psychological test constructor or teacher, what precautions are to be considered, while preparing a question paper or test paper.

1.14 SUGGESTED READINGS

Aggarwal, Y.P.: *"Statistical Methods-Concepts, Applications and Computation"*. New Delhi: Sterling Publishers Pvt. Ltd.

Ferguson, G.A.: *"Statistical Analysis in Psychology and Education"*. New York: McGraw Hill Co.

Garrett, H.E. & Woodworth, R.S.: *"Statistics in Psychology and Education"*. Bombay : Vakils, Feffer & Simons Pvt. Ltd.

Guilford, J.P. & Benjamin, F.: *"Fundamental Statistics in Psychology and Education"*. New York: McGraw Hill Co.

Srivastava, A.B.L. & Sharma, K.K.: *"Elementary Statistics in Psychology and Education"*, New Delhi: Sterling Publishers Pvt. Ltd.

UNIT 2 SIGNIFICANCE OF MEAN DIFFERENCES, STANDARD ERROR OF THE MEAN

Structure

2.0 INTRODUCTION

The main function of statistical analysis in behavioural sciences is to draw inferences or made generalisation regarding the population on the basis of results obtained. Therefore the infrential statistics is that branch of statistics which primarily deals with inferences from a sample to a large population from which the sample has been taken. This depends on the fact that how good is the sample estimate. If the sample estimate is not good i.e. having the considerable error or not reliable, we could not be able to draw correct or good inference about the parent population. Thus before to draw inference about the whole population or to made generalisation, it is essential first to determine the reliability or trustworthiness of computed sample mean or other descriptive statistical measures obtained on the basis of a sample taken from a large population.

As an implication of the trustworthiness of the sample measures, we are concerned also with the comparison of two sample estimates with a view to find out if they come from the same population. In other words, the two sample estimates of a given trail of the population do not differ significantly from each other.

Here significant difference means a difference larger than expected by chance or due to sampling fluctuations.

Thus the present unit, highlights the concept of standard error of a sample mean and to compare the two sample means drawn randomly from large population. So that we may be able to test our null hypothesis scientifically, which is made in relation to our experiment or study and to draw the inferences about the population authentically.

2.1 OBJECTIVES

After going through this unit, you will be able to:

- Define and explain the meaning of inference;
- Describe the concept of statistics and parameters;
- Distinguish between statistics and parameters;
- Explain the meaning of significance, significance level;
- Elucidate their role and importance to draw inference and to make generalisation about the population;
- Explain and differentiate between Sampling Error, Measurement Error and Standard;
- Error of Mean value obtained on the basis of a sample from a population;
- Analyse the 't' distribution and its role in inferential statistics;
- Describe the standard error of large and small size sample means;
- Analyse the mean of the population on the basis of the mean of a sample taken from the population with certain level of confidence;
- Determine the appropriate sample size for a experimental study or for a research work Compare the means of two sample means obtained from the same population;
- Differentiate between independent sample means and correlated sample means;
- Test the null hypothesis (Ho) made in relation to an experimental study; and
- Analyse the errors made in relation to testing the null hypothesis.

2.2 THE CONCEPT OF PARAMETERS AND STATISTICS AND THEIR SYMBOLIC REPRESENTATION

Suppose you have administered a verbal test of intelligence on a group of 50 students studying in class VIII of a school of your city. Further, suppose you find the mean I.Q. of this specified group is "105". Can you from this data or information obtained on the relatively small group, say any thing about the I.Q. of all the VIII class students studying in your city. The answer is "Yes" but under certain conditions. The specified condition is "the degree to which sample mean (M) which is also known as "Estimate" represents its parent population mean which is known as "True Mean" or "Parameter". Therefore the two terms Estimates and Parameters are defined as given below.

2.2.1 Estimate

The statistical measurements e.g. measures of central tendency, measures of variations, and measures of relationships obtained on the basis of a sample are known as "Estimates" or Statistics. Symbolically, these are generally represented by using the English alphabets e.g.

Mean = M, Standard Deviation = S.D. or σ, Correlation = r etc.

2.2.2 Parameter

The statistical measurements obtained on the basis of entire population are known as "True Measures" or "Parameters".

Symbolically, these are represented by putting over the bar (-) over corresponding English alphabets or represented by Greek letters e.g.

True Mean or Population Mean = $\overline{M}$ or μ (Mu)

True S.D. or Population S.D. = $\overline{S.D.}$ or $\bar{\sigma}$

True or Population correlation = $\bar{r}$ or η

It is rarely if ever possible to measure all the units or members of a given population. Therefore, practically or for case we draw a small segment of the population with convenient specified number of units or members, which is known as the sample of the population.

Therefore, we do not know the parameters of a given population. But we can under specified condition, forecast the parameters from our sample statistics or estimates with known degree of accuracy.

2.3 SIGNIFICANCE AND LEVEL OF SIGNIFICANCE OF THE STATISTICS

Ordinarily, we draw only a single sample from its parent population. However, our problem becomes one of determining how we can infer or estimate the mean of the population (Mpop) on the basis of the sample mean (M). Thus the degree to which a sample mean (M) represents its parameter is an index of the "Significance" or Trustworthiness of the computed sample mean.

When we draw a sample from the population, the observed statistics or estimate that is the mean of the sample obtained, may be some time large or small to the mean of the population (Mpop). The difference may have arisen "by chance" due to the differences in the composition of our sample, or due to its selection method or the procedure followed in the sample selection. The gap between the two measures sample mean (M) and population mean (Mpop), if is low and negligible the sample mean is considered to be trustworthy and we can forecast or estimate the population mean (Mpop) successfully.

Therefore, a sample mean (M) is statistically trustworthy or significant to forecast the mean of the population (Mpop), depending upon the probability that the difference between the two measures i.e. Mpop and M could have been arisen "by chance".

And the confidence level to which this forecast has been made is known as level of confidence or level of significance.

In simpler terms, the level of significance or level of confidence is a degree to which we accept or reject or predict a happening or incidence with confidence.

There are a number of levels of confidence or levels of significance e.g. 100%, 99%, 95%, 90% .... 50% etc. In psychology and other behavioural sciences, generally, we consider only two levels of significance viz. the 99% level of significance or level of confidence and 95% level of significance a level of confidence.

The amount 99% and 95% confidence is also termed as 0.01 and 0.05 level of confidence. The 0.01 level means, if we repeatedly draw a sample or conduct an experiment 100 times, only on one occasion, the obtained sample mean or results will fallout side the limits Mpop ± 2.58 S.E.

Here the term S.E. means the standard error exists in the estimate or sample statistics.

Similarly 0.05 level means, if repeatedly draw a sample or conduct an experiment 100 times, only on five occasions the obtained sample mean will fall out side the limits Mpop ± 1.96 S.E.

The value 1.96 and 2.58 have been taken from the 't' distribution or 't' table (P...).

Keeping large size sample in view.

The 0.01 level is more rigorous and higher in terms of standard, as compared to the 0.05 level and would require a high level of accuracy and precision. Hence, if an obtained value (on the basis of a sample or an experiment) is significant at 0.01 level, it is automatically significant at 0.05 level but the reverse is not always true.

2.4 SAMPLING ERROR AND STANDARD ERROR

The score of an individual of a group obtained on a certain test consists of two types of errors (i) measurement error and (ii) statistics error.

In other words,

True Score X_T = observed score or obtained score (X_0) ± error (E)

and error E is = Measurement Error + Statistics Error

The measurement error is caused by a measuring instrument used to measure a trait or variable and personal observation made by the individual on the instrument.

The measurement error is due to the reliability of a test, as no test is perfectly reliable specially in behavioural sciences. In other words no test or a measuring instrument gives us 100% accurate measurement. The personal error is dependent upon the accurate perception and attention of the individual to take observations or measurement on the measuring instrument.

The statistics error refers to the errors of sample statistical measurements or estimates obtained on the basis of a sample drawn from a population.

As it is not possible to have perfectly true representative sample of a population in behaviourable sciences. The statistics error is of two types: (i) Sampling Error (ii) Standard Error of statistics i.e. statistical measurements. Now let us see what are these two errors in detail.

2.4.1 Sampling Errors

Sampling error refers to the difference between the mean of the entire population and the mean obtained of the sample taken from the population.

Thus sampling Error = Mpop ~ M or $\overline{M}$ - M

As the difference is low the mean obtained on the basis of sample is near to the population mean and sample mean is considered to be representing the population mean (or Mpop)

2.4.2 Standard Error

The standard error is nothing but the intra differences in the sample measurements of number of samples taken from a single population.

As the intra differences in the number of sampling observation i.e. the statistics of the same parent population is less and tending to zero, we may say the obtained sample statistics is quite reliable and can be considered as representative of the Mpop or $\overline{M}$.

For more clarification, suppose you wish to determine the I.Q. level of the high school going students studying in the various schools of your district. Is it possible for you to administer the intelligence test on all the high school going students of your district and get the Average I.Q.? Your answer may be certainly not.

The easiest method is to select a sample of 10 schools each from urban and rural areas of your district by using random method and administer the intelligence test to the students studying in high school class of these selected 20 schools in total. In such condition, you may have approximately 20 samples of high school going students and have 20 means of intelligence scores obtained on the intelligence test which you have administered. It is possible that all the mean values you have obtained are not equal. Some may be small and some may be large in their values. Theoretically, there should be no difference in the mean value obtained and all the mean values should be equal as all the samples are taken from the same parent population by using random method of sample selection. Thus the inter variation lies within the values of 20 mean values indicating that the error lies within the various observations taken.

Further, to have the mean I.Q. of all the high school going students you may calculate combined mean or average mean of all 20 means obtained on the basis of samples taken. This obtained combined or average mean value is the I.Q. of the parent population i.e. the high school going students of your district. If you compare all the

20 sample means to the obtained combined mean value i.e. the population mean, you will find that some of sample means are lesser than this Mpop value, and some are higher.

Further one step more, you calculate the difference of these sample means from the population mean obtained, i.e. find (Mpop – M_1), (Mpop – M_2) .... (Mpop – M_{20}). You will find that some of these difference values are negative and some are positive. The mean of all these differences should be zero and the standard deviation of all these differences should be 1.

As the errors are normally distributed in the universe, in simple terms we can say that we have a normal distribution of specific statistics or sample statistical measurement, which is also known as sampling distribution.

Therefore, theoretically the standard error of the statistics (sample statistical measurements) is the standard deviation of the sampling distribution of the statistics and is represented by the symbol $S.E._M$.

Standard Error of sampling measurements on statistics is calculated by using the formula given below :

$$\text{S.E.M or } \sigma_M = \frac{\bar{\sigma}}{\sqrt{N}}$$

Where, S.E.M = Standard Error of sampling measurement

$\bar{\sigma}$ = Standard deviation of the scores obtained from the population.

N = Size of the sample or total number of units in a sample

Look carefully and study the formula given above, you will find, the standard error of any statistics depends mathematically upon two characteristics.

i) the variability or spread of scores around the mean of the population and

ii) the number of units or cases in the sample taken from the population.

As there is low variability in the scores of population, i.e. the population is homogeneous on the trait being measured, and also the number of cases in the sample are too large, the standard error of the statistics is tending to zero.

In the formula, standard error of statistics is directly proportionate to the standard deviation (σ) of the scores of population and inversely proportionate to the size of sample or number of cases in the sample (N).

Thus in brief, it can be said that if the population is homogeneous to the variable or trait being measured and a large size of sample (say more than the 500 units), taken from the population; in such condition the sample drawn will be representative to its parent population and is highly reliable.

Self Assessment Questions

1) Explain the following terms:

a) Estimate

....

....

b) Parameter

....

....

c) Statistics

....

....

d) Sampling Error

....

....

e) Measurement Error

....

....

f) Standard Error

....

....

2) What is the general formula to know the standard errors of the various statistical measures?

....

....

....

3) What do you mean by significance and levels of significance?

....

....

....

4) In behavioural sciences, which levels of confidence are considered

....

....

....

5) What is the difference between significance of statistics and confidence interval for true statistics?

....

....

....

2.5 'T' RATIO AND 'T' RATIO DISTRIBUTION TABLE

2.5.1 't' Ratio

So far we have studied two types of distributions viz.,

1) Distribution of scores – Normal distribution (unit I)
2) Distribution of 'statistics' or sample statistical measures (p -)

For more clarification, again going back. suppose we have drawn 100 samples of equal size (say n = 500) from a parent population and we calculate the mean value of the scores of a trait of the population obtained from each sample. Thus we have a distribution of 100 means.

Of course all these sample means will not be alike. Some may have comparatively large values and some may have small. If we draw a frequency polygon of the mean values or the "statistics" obtained, the curve will be "Bell – Shaped" i.e. the normal curve and having the same characteristics or properties as the normal probability curve has.

The distribution of statistics values or sample statistical measurements is known as the "sampling – distribution" of the statistics.

The corresponding standard score formula i.e. $z = \frac{X-M}{\sigma}$, may now become as –

$$t = \frac{M-M_{pop}}{S.E._M} \quad \text{or} \quad \frac{M-\overline{M}}{S.E.\overline{M}}$$

where t = standard score of the sample measures or statistics and termed as "t.Ratio"

M = Mean of specific statistics or sample measure.

$\overline{M}$ or Mpop = Mean of the parameter value of the specific statistics or mean of the specific statistics of the population

$S.E._M$ = Standard Error of the statistics i.e. the standard deviation of the sampling distribution of the statistics.

Actually, t is defined as we have defined the z. It is the ratio of deviation from the mean or other parameter, in a distribution of sample statistics, to the standard error of that distribution.

To distinguish z score of the sampling distribution of sample statistics, we use "t" which is also known as "student's t".

The "t" ratio was discovered by an English statistician, W.S. Gossett in 1908 under the pen name "student". Therefore, the "t" ratio is also know as "student's t" and its distribution is known as "student's t distribution".

As the "t" ratio is the standard score (like z score) with mean = 0 and standard deviation ±1, therefore the t ratio is a deviation of sample mean (M) from population mean ($\overline{M}$ or Mpop).

If this deviation is large the sample statistics mean is not reliable or trustworthy and if the deviation is small, the sample statistics mean is reliable and representative to the mean of its parent population ($\overline{M}$).

2.5.2 The Sampling Distribution of "t" Distribution

Just now we have studied about the sampling distribution of sample statistics and the "t" ratio. Imagine that we have taken number of independent samples with equal size from a population. Let us say we have computed the "t" ratio for every sample statistics with N constant. Thus a frequency distribution of these ratios would be a sampling distribution of "t" and is known as "t" distribution. The mean of all "t" ratios is zero and standard deviation of all "t" ratios i.e. σ is always $\pm\sigma_t$.

It has been observed that if the sample size varies the sampling distribution of "t" also varies, though it is normal distribution. The sampling distribution of "t" may vary in kurtosis. Student's t distribution becomes increasingly leptokurtic as the size of sample decreases.

As the size of sample is tending to be large, the distribution of "t" approaches to the normal distribution. Thus we have a family of "t" distributions, rather to one and the σt values varies on the x axis.

Fisher has prepared a table of "t" distribution having N, i.e. the size of sample different for different levels of significance. The details of the same are given below:

Table 2.5.1 : Table of "t" for use in determining the significance of statistics

Degrees of Freedom	Probability (P) 0.10	0.05	0.02	0.01
1	$t = 6.34$	$t = 12.71$	$t = 31.82$	$t = 63.66$
2	2.92	4.30	6.96	9.92
3	2.35	3.18	4.54	5.84
4	2.13	2.78	3.75	4.60
5	2.02	2.57	3.36	4.03
6	1.94	2.45	3.14	3.71
7	1.90	2.36	3.00	3.50
8	1.86	2.31	2.90	3.36
9	1.83	2.26	2.82	3.25
10	1.81	2.23	2.76	3.17
11	1.80	2.20	2.72	3.11
12	1.78	2.18	2.68	3.06
13	1.77	2.16	2.65	3.01
14	1.76	2.14	2.62	2.98
15	1.75	2.13	2.60	2.95
16	1.75	2.12	2.58	2.92
17	1.74	2.11	2.57	2.90
18	1.73	2.10	2.55	2.88
19	1.73	2.09	2.54	2.86
20	1.72	2.09	2.53	2.84
21	1.72	2.08	2.52	2.83
22	1.72	2.07	2.51	2.82
23	1.71	2.07	2.50	2.81
24	1.71	2.06	2.49	2.80
25	1.71	2.06	2.48	2.79
26	1.71	2.06	2.48	2.78
27	1.70	2.05	2.47	2.77
28	1.70	2.05	2.47	2.76
29	1.70	2.04	2.46	2.76
30	1.70	2.04	2.46	2.75
35	1.69	2.03	2.44	2.72
40	1.68	2.02	2.42	2.71
45	1.68	2.02	2.41	2.69
50	1.68	2.01	2.40	2.68
60	1.67	2.00	2.39	2.66
70	1.67	2.00	2.38	2.65
80	1.66	1.99	2.38	2.64
90	1.66	1.99	2.37	2.63
100	1.66	1.98	2.36	2.63
125	1.66	1.98	2.36	2.62
150	1.66	1.98	2.35	2.61
200	1.65	1.97	2.35	2.60
300	1.65	1.97	2.34	2.59
400	1.65	1.97	2.34	2.59
500	1.65	1.96	2.33	2.59
1000	1.65	1.96	2.33	2.58
∞	1.65	1.96	2.33	2.58

Let us now take an example. Let us say there are 26 subjects. N = 26.

Example: When N = 26, the corresponding degree of freedom (df) is N-1 i.e. 25.

In column 3 at 0.05 level of significance the t value is 2.06.

It means that five times in 100 trials a divergence of sample mean or statistics obtained may be expected at a 2.05 σ to M, that is to its mean population either to its left or right side.

The "t" distribution table has great significance in inferential statistics testing the null hypothesis framed in relation to various experiments made in psychology and education.

2.6 STANDARD ERROR OF SAMPLE STATISTICS – THE SAMPLE MEAN

The standard error of sample statistics or the statistical measurements of a sample has great importance in inferential statistics. With the help of the standard error statistics we can determine the reliability or trustworthiness of the descriptive statistics e.g. proportion percentage, measures of central tendency (mean, median and mode) measures of variability (standard deviation & quartile deviation), Measures of correlation (r, p and R) etc.

For convenience here we discuss only the significance of means which are detained as under :

2.6.1 Meaning of Standard Error of Mean

The Standard error of mean measures the degree to which the mean is affected by the errors of measurement as well as by the errors of Sampling or Sampling fluctuations from one random sample to the other. In other words how dependable is the mean obtained from a sample to its parameter i.e. population Mean (M pop).

Keeping in mind the Sample Size; there are two situations:

i) Large Sample

ii) Small Sample

2.6.2 The Standard Error of Mean of Large Sample

When we say large sample, the number of items in the sample will be more than 30. That is N>30. In such condition the Standard error of the Mean is determined by using the formula given below:

$S.E._M = \sigma / \sqrt{N}$ when N= >30

Where

$S.E._M$= Standard Error of the Mean of the scores of a large sample

σ = Standard deviation of the scores of a population

N = Size of the sample or number of cases in the sample

This formula is used when the population parameter of standard deviation (σ) is known. But in practice it is not possible to have the value of Ïó. Having the situation that the sample is selected from the population by using random method of sample selection, σ can be replaced by the ó. The value of the standard deviation of the scores of the sample taken. Therefore, in the above formula the σ can be replaced by ó.

S.E.$_M$ = ó / $\sqrt{N}$ when N >30

where

S.E.$_M$ = Standard Error of the Mean of the scores of a Sample

ó = Standard Deviation of the scores of a sample.

N = Number of units or cases in the Sample.

Example 1: A reasoning test was administered on a sample of 225 boys of age group 15 + years. The mean of the scores obtained on the test is 40 and the standard deviation is 12. Determine how dependable the mean of sample is.

Given : N=225. M=40 and ó = 12

To find : The trustworthyness of the sample mean we know that standard error of the mean, when N>30 is determined by using the formula-

S.E.$_M$ = ó / $\sqrt{N}$

S.E.$_M$ = 12 / $\sqrt{225}$

= 12 / 15 = 0.80

Or S.E$_M$ = 0.80

i.e. = 0.80

Interpretation of the Result

Keeping in mind the logic of sampling distribution, that is if we draw 100 samples, each sample has 225 units from a large population of boys of age group 15+ years, we will have 100 sample means falling into a normal distribution around the Mpop and ó$_M$ (the standard deviation of sampling distribution of Means i.e. the standard error of Mean)

As per properties of Normal Distribution, in 95% cases the sample means will lie within the range of ±1.96 in to the Mpop (see Z table in unit I). Conversely out of 100, the 99 sample means having equal size, will be within the range of ±2.57 (2.57 ×0.80) of the Mpop.

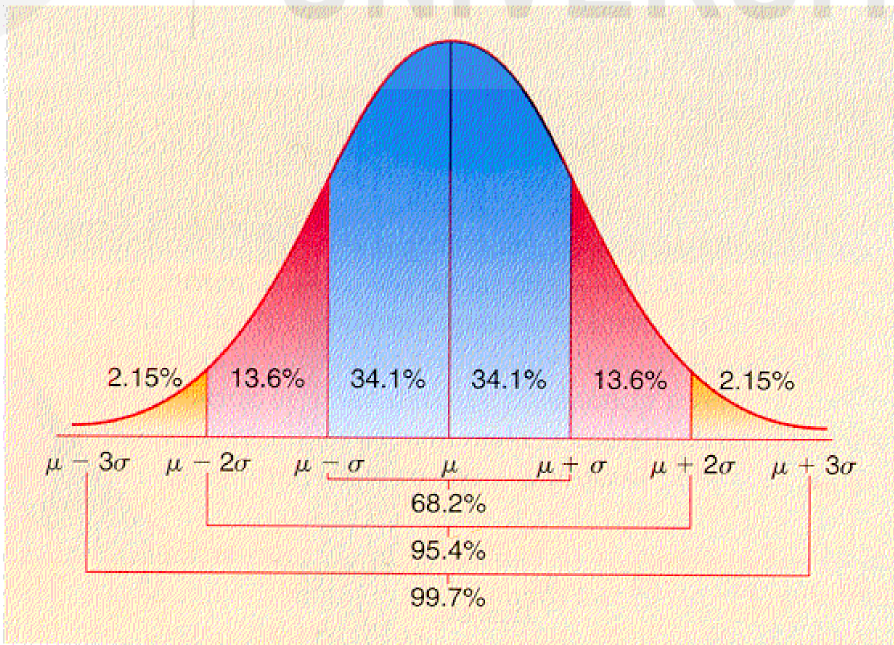

Fig. 2.6.1: Image graph

(**Source:** kmblog.rmutp.ac.th/.../28/normal-distribution/)

From the figure it is quite evident that the S.E. of the M=40. Sample of 225 having

bó = 12 lie within the acceptable region of the N.P.C.(Normal Probability curve). Thus the sample mean obtained is quite trustworthy with the confidence of 95% probability. There are only 5% chances that the sample mean obtained will lie in the area of the rejection of M.P.C.

In simplest term we can say that, there is 95% probability the maximum possibility of the standard error of the sample mean (40) is ± 1.57 (1.96×0.80) which is less than the value of T=1.96 at .05 level of confidence for df=224 (N-1) Thus the obtained sample mean (40) is quite dependable to its Mpop with the confidence level of 95%.

Example 2: In the example 1, suppose in place of N=225, we have a sample of 625 units and the remaining observations are the same. Determine how good an estimate is it of the population mean?

Solution

Given : N=625, M=40 and ó=12

To find : Dependency of sample Mean or reliability of sample mean

We know that

$$ó_M \,/\, S.E._M = \frac{\sigma}{\sqrt{N}}$$

$$= \frac{12}{\sqrt{625}}$$

$$= \frac{12}{25} = 0.48$$

Or $S.E._M$=0.48.

Interpretation of Result

The maximum standard error of sample M=40 and Σ =12 having 625 units is ±0.94 (1.96×0.48) at 95% level of confidence which is much less than the value of t.05 = ± 1.96. Therefore, the obtained sample mean is reliable and to the considered as representative to its Mpop at 95% level of confidence.

Self Assessment Questions

1) Compare the two results obtained from Example no 1 and 2 respectively. What you have observed and what is your conclusion.

....

....

....

....

....

2) The mean achievement score of a random sample of 400 psychology students is 57 and D.D. is 15? Determine how dependable is the sample mean?

....

....

....

....

3) A sample of 80 subjects has the mean = 21.40 and standard deviation 4.90. Determine how far the sample mean is trustworthy to its Mpop.

....

....

....

....

2.6.3 Degree of Freedom

Before to proceed for the standard error of small sample mean, it is imperative here to understand the concept of Degrees of Freedom.

The expression degrees of freedom is abbreviated from the full expression "Degrees of Freedom to Vary". A point in the space has unlimited freedom to move in any direction, but a point on a straight line has only one freedom to move on the line i.e. it is free to move in one dimension only. A point on the plane has two degrees of freedom. A point on a three dimension space having three degree of freedom to move.

This shows that the degree of freedom is associated to the number of restrictions imposed upon the observations.

The degree of freedom is a mathematical concept and is a of key importance in inferential statistics. Almost all test of significance require the calculation of degree of freedom.

When a sample statistics is used to estimate a parameter, the number of degrees of freedom depends upon the number of restrictions placed upon the scores, each restriction reducing one degree of freedom (df).

For example, we have four numbers 4, 5, 8 and 3, the sum of these numbers is 20 which is fixed. e.g. here we have restricted freedom to get sum 20, as we have only to change the values of first three figures and not the last one, as it depends upon the fixed sum 20,

i.e.

6 + 6 + 9 +———-20

Or 2 + 7 + 10 + ———20

Or 6 + 7 + 5 +———20

In the above sum expressions we have last one freedom to determine the value of last 4th numeral figure as to get fixed sum 20. Therefore in the above expressions, we are bound to take forcefully the numeral figures 4,1 and 2 respectively.

Like wise, in Statistics, when we calculate Mean or S.D. of a given distribution of scores we lose one degree of freedom to get fixed sums. Therefore we have N-1 df to compute statistical measures , specifically the standard deviation of the scores given and N-2 in case to compute the co-efficient of correlation. It means the degree of freedom is not always (N-1) however, but will vary with the problem and the restrictions imposed.

In the case of very large sample used in behavioural sciences or social sciences no appreciable difference takes place in the value of $ó_M$ by N-1 instead of N. The use of N or N-1 thus remains a matter of arbitrary decision. But in the case of small samples having number of units or cases below 30 $ó_M$ no correction (N-1) has been applied in computation of S.D and thus a considerable variation occurs on the value of $ó_M$ … Therefore, it is imperative to use N-1 in place of N in computation of $ó_M$ of the small sample.

Further you have to study the t" distribution table (table No-2.5.1) very carefully. In the process, you will find that as the size of sample or degree of freedom approaches to 500 or above, the "t" value approaches to the value of 95% and 99% are 1.96 and 2.58 respectively and remain constant. It means the "t" distribution becomes normal distribution or Z-distribution. When the size of sample decreases especially below 30 you will find the "t" values are gradually increasing at 95% level and 99% level considerably. In such condition the $ó_M$ values gives us wrong information and we may interpret the results inappropriately.

2.6.4 The Standard Error of Means of Small Sample

The small sample means, when size of sample (N) is about 30 or less, is treated as small sample. The formula for the standard error of small sample mean score is as follows–

$$S.E._M \text{ on } S_M = ó \,/\, \sqrt{N-1}$$

As here $S.E_M$= Standard error of Mean of Small sample

Standard deviation of the population

N = Size of the sample i.e. 30 or below

Note : For practice we replace -ó by ó i.e. standard deviation of the sample. Because of the reason ó is not possible to obtain for whole population.

Example 3: A randomly selected group of 17 students were given a word cancellation test. The mean and S.D obtained for cancelling the words per minute is 58 and 8 respectively. Determine how far sample mean is acceptable to represent the Mean of the population?

Solution

Given : N=17, M=58 and ó = 8

To find : dependency of the sample mean

In the problem the size of sample is less than 30. Therefore to find the standard error of sample mean is

$$S.E_M = \frac{\sigma}{\sqrt{N-1}}$$

$$S.E_M = \frac{8}{\sqrt{17-1}}$$

$$= \frac{8}{\sqrt{16}}$$

$$= 8 / 4 = 2$$

In the "t" table (table no. . 2.5.1) at .01 level, the value of "t" for 16 df is 2.92 and the obtained value of t = 2.00. which is less in comparison to the "t" value given in the able. Therefore, the obtained sample mean (58) is quite trustworthy and representing its Mean population by 99% confidence. There is only one chance out of 100, that sample mean is low or high.

Self Assessment Questions

1) What is the concept of Degree of Freedom (df)?

....

....

....

....

2) Why we consider the (df) which determining the reliability or trustworthiness of the statistics.

....

....

....

....

3) What is the difference to calculate the standard error of Mean of Large Size and Small Size samples.

....

....

....

....

2.7 APPLICATION OF THE STANDARD ERROR OF MEAN

2.7.1 Estimation of the Population Statistics – The Mpop

The wider use or application of $S.E._M$ is to estimate the population Statistical measurements i.e. the M pop. Here we are concerned only with the mean. Therefore we will discuss to estimate the mean of population on the basis of standard error of the mean obtained on either large size sample or small size sample.

Here, it is important to note that the estimation is always in range rather to point estimation because exact single value of any measurement is not possible. For example a student can not forecast with confidence that he will secure 85 to 95 marks out of 100 in statistics in the final examination.

Therefore, estimation of Mpop is always in range rather to point. Thus the limits obtained of Mpop (The lower and upper limits) are also known as Fiduciary limits.

The term Fiduciary limits was used by R.A. Fisher for the confidence interval of parameter and the confidence placed in the interval defined as Fiduciary Probability.

The simplest formula to estimate the Mpop is as under:

Mpop or M = M± 2.58 σ_M ——— (at .01 level of significance)

Mpop or M= M±1.96 σ_M ————(at .05 level of significance)

For more classification study the following examples carefully:-

Example 4: One language test was given to 400 boys of VIII class, the mean of their performance is 56 and the standard deviation is 14. What will be the Mean of the population of 99% level of confidence?

Solution

Given: N=400, M=56 and σ =14

To find out: Estimation of population mean at 99% level of confidence.

We know that Mpop at .01 level or at 99% confidence level is

Mpop .01 or M.01= M±2.58 σ_M

Where

M .01: Mean of the population at 99% confidence level of confidence

M : Mean of the sample

σ_M : Standard Error of the sample mean.

In the problem the values of Mean and N are known and the value of σ_M is unknown. The value of σ_M can be determined by using the formula.

$$\sigma_M = \frac{\sigma}{\sqrt{N}}$$

$$\therefore \sigma_M = \frac{14}{\sqrt{400}} = \frac{14}{\sqrt{20}}$$

Or σ_M = 0.70

Thus $\overline{M}$.01 = 56± 2.58×0.70

= 56±1.806

= 54.194-57.806

Or $\overline{M}$.01 = 54-58

The Mean of the population at 99% level of confidence will be within the limits 54 to 58. In other words there are 99% chances that the Mean of the population lie within the range 54-58 scores. There is only 1% chance that mean of the population lie beyond this limit.

Example 5: A randomly selected group of 26 VI grade students having a weight of 35 kg and S.D = 10 kg. How well does this value estimate the average weight of all VI grade students at .99 and .95 level of confidence?

Solution

Given : N =26, M = 35kg and ó= 10 kg.

To find out : The fiduciary limits of the population mean at .05 and .01 levels.

In the problem the given sample size is below 30, Therefore to have the standard error of sample mean we will use the formula:

S_M = 2.0

And = N – 1 = 26 – 1

Or df = 25

i) Fiduciary limits of M at .01 level of confidence

By consulting the t table, level of confidence, the value of "t" for 25 df is 2.79

Thus, the Fiduciary limit of M at .01 or 99% level is

$= \bar{M} \pm 2.79\sigma_M$

$= 3.5\pm 2.79\times2.00$

$= 3.5\pm5.50$

$\therefore \bar{M}.01 =$ 29.42-38.58 = (9.16)

ii) The fiduciary limits of $\bar{M}$ at .05 level of confidence

$M_{0.05} = M\pm 2.06\ \sigma_M$

$= 35\pm2.06\times2.0$

$= 35\pm4.12$

Or $M_{0.05} =$ 30.88–39.12

i) Thus The Fiduciary Limits of M .01 = 29.42-38.58

ii) The Fiduciary Limits of M.05 = 30.88-39.12

2.7.2 Determination of the Size of Sample

Standard error of the statistics is also used to estimate the sample size for test results. In order to learn how to use the standard error of the statistics you must go through the following examples.

Example 6

If the standard deviation of a certain population (ó) is 20. How many cases would require in a sample in order that standard error of the mean should not miss by 2.

Solution

Given: ó= 20, and S. $E._M$=20.

To Find Out : No of cases in the sample to be selected i.e. to determine the Size of Sample (N)

We know that

$$\text{S.EM} = \frac{\sigma}{\sqrt{N}}$$

$$\therefore \; 2 = \frac{20}{\sqrt{N}}$$

$$\text{Or } \sqrt{N} = \frac{20}{2} = 10$$

Or $N = (10)^2$

Or $N = 100$

If the standard error of the sample mean should not be more than 2 in such condition the maximum sample Size Should be 100 i.e. N=100 .

Example 7: The standard deviation of the intelligence scores of an adolescent population is 16. If the maximum acceptable standard error of the mean of the sample should not miss by 1.90, what should be the best sample size at 99% level of confidence?

Solution

Given : ó = 16, SE_M = 1.90

To find out : Sample size which represent its parent population upto the level of 99%.

We know that the Z value of 99% cases is 2.58 (From Z Table)

It means due to chance factors the sample mean would deviate from Mpop by 2.58 $ó_{M.}$ Further in keeping view the measurement and other uncontrolled factors, the measured error in the sample mean we would like to accept is 1.90.

Therefore the maximum error in the sample which we would like to select from the parent population is

$$\text{S.E}_M = ó \times \frac{2.58}{\sqrt{N}}$$

$$\text{Or } \sqrt{N} = \frac{\sigma \times 2.58}{\text{SE}_M}$$

$$\text{Or } N = \left(\frac{\sigma \times 2.58}{\text{SE}_M}\right)^2$$

$$\therefore N = \left(\frac{16 \times 2.58}{1.90}\right)^2$$

Or N = 472

To have a representative sample up to the level of 99% to the parent population, it is good to have a sample size more than 472 cases.

Self Assessment Questions

1) Given M =26.40, ó = 5.20 and N=100 compute

   The fiduciary limits of True Mean at 99% confidence interval

   The fiduciary limits of Population Mean at .95confidence interval.

   ....

   ....

   ....

   ....

2) The mean of 16 independent observations of a certain magnitude is 100 and S.D is 24. At .05 confidence level what are the fiduciary limits of the True Mean.

   ....

   ....

   ....

   ....

3) Suppose it is known that S.D of the scores in a certain population is 20. How many cases would we in a sample in order that the S.E of the sample mean be σ2.

   ....

   ....

   ....

   ....

2.8 IMPORTANCE AND APPLICATION OF STANDARD ERROR OF MEAN

The Standard error of statistics has wide use in inferential statistics. It helps the experimenter or researcher in drawing concrete conclusions rather than abstract ones.

The various uses of standard error of the statistics are as under:

Various devices are used for determining the reliability of a sample taken from the large population. The reliability of the sample depends upon the reliability of the statistics, which is very easy to calculate.

The main focus of the standard error of statistics is to estimate the population parameters. No sampling device can ensure that the sample selected from a population may be representative. Thus the formula of the standard error of statistics provides us the limits of the parameters, which may remain in an interval of the prefixed confidence interval.

The method of estimating the population parameters the research work feasible, where the population is unknown as impossible to measure. It makes the research work economical from the point of crew of time. Energy and money.

Another application of the standard error of the statistics is to determine the size of the sample for experimental study or a survey study.

The last application of the standard error of statistics to determine the significance of difference of two groups to ascertained by eliminating the sampling or change by estimating the sampling or change errors.

2.9 THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN TWO MEANS

Suppose we wish to study the linguistic ability of the two groups say boys and girls of age group 15 years.

First we have to select two large and representative samples from the two different populations. The population of the Boys of age group 15 years old and second, the population of girls of same age as of the Boys has has to be selected.

To administer the linguistic ability test Battery to both the groups of Boys and Girls selected as sample.

To compute the mean values of like scores obtained by the two groups on the linguistic ability test battery and find the difference between them.

After the above procedural steps, suppose that there is a difference in the means of the two groups and the difference is in favour of the girls of age group 15+ years old. Is this evidence sufficient to draw the conclusion that girls are superior in linguistic ability in comparison to the boys having same age level?

Probably the answer to this question may either be "yes" or "No" depending upon further testing of the difference of the means of two groups whether it is statistically significant or not. In other words, it is essential to test further, that how far the difference exists in the mean values of two groups is due to " chance" factor or is it "real and dependable"

This question involves the standard error of the difference that exists between the mean values of the two groups and the same as significant or not. Therefore, in order to test the significance of an obtained difference, we must first have a S.E. of the difference of sample means. Then from the difference between the sample means and the standard error of the difference of sample means (S.E. $_{D.M.}$), we can determine whether a difference probably exists between the population means.

A difference is called *Significant* when the probability is high that it cannot be attributed to chance (i.e. by temporary and accidental factors or by sampling fluctuations) and hence represents a true difference between population means. And a difference is *non-significant* or chance, when it appears reasonably certain that it could be easily have arisen from sampling fluctuations and hence implies no real or true difference between the population means.

Thus the above discussion leads us to conclude that the significance of the difference between two sample means obtained from the two populations either independent or correlated depend upon two factors, viz.,

i) Standard Error of Difference between the two means, and

ii) The levels at which S.E $_{DM}$ is significant.

iii) Standard Error of the difference of two means (Σ_{DM}) and critical ratio (C.R)

2.9.1 Standard Error of the Difference of Two Means and Critical Ratio (CR)

Suppose we have two independent large populations, say A and B, and let us say that we have taken several numbers of samples (say two) from each population. Now if we compute the mean values of the scores of a trait of the two populations, we have 100 sample means obtained from the population A and 100 sample means obtained from the population B, and let us say that we find that there is a difference between the two sample means of population A and B. Thus in this way we have 100 differences of sample means. If we plot the frequency polygon of these hundred samples, certainly we will have a normal curve, and the distribution of the sample mean differences will be known as sampling *Distribution of Mean Differences.*

The standard error of the sample mean differences can be obtained by computing standard deviation of the sampling distribution of mean differences. This can be computed by using the formula:

$$S.E_M \text{ or } \Sigma_{DM} = \frac{\Sigma 1^2 + \Sigma 2^2}{N_1 + N_2} \text{ (In case of two independent population)}$$

Where

Σ_1 = Standard Deviation of the scores of a trait of the sample-1

Σ_2 = Standard Deviation of the scores of a trait of the sample-2

N_1= Number of cases in sample-1

N_2= Number of cases in sample-2

After having the standard error of the sample mean differences, the next step is to decide how far the particular sample mean difference is deviating from the two population mean differences ($M_1 \sim M_2$) on the normal probability curve scale. For the purpose we have to calculate Z score of the particular two sample mean differences, using the formula

$$Z = \frac{X–M}{\sigma_{DM}} \quad \text{(see unit-I)}$$

or

$$Z = \frac{(M_1 \sim M_2)–(M_1 \sim M_2)}{\frac{\sqrt{\Sigma 1^2 + \sigma^2}}{N_1+N_2}}$$

To distinguish the Z score of the difference of two sample means, the symbol C.R (Critical Ratio) is used. Therefore

$$C.R = \frac{(M_1 \sim M_2)–(M_1 \sim M_2)}{\Sigma_{DM}}$$

If the two independent populations are alike or same about a trait being measured, then

$$M_1 \sim M_2 \quad = 0$$

$$\therefore \text{C.R.} = \frac{(M_1 \sim M_2) - 0}{\sigma_{DM}}$$

$$\text{Or} \quad \text{C.R} = \frac{(M_1 \sim M_2)}{\sigma_{DM}}$$

This is the general formula to decide the significance of the difference exists in the two sample means taken from the two independent populations.

The formula of C.R. clearly indicates that it is a simple ratio between difference of the two sample means and the standard error of the sample mean differences. Further it is nothing but a Z score, which indicates how far the two sample mean difference is deviating from the two parent population mean difference, which is Zero.

2.9.2 Levels of Significance

Whether a difference in the two statistic i.e. the statistical measures obtained or the parameters are to be considered as statistically significant?

It depends upon the probability that the given difference could have arisen "by chance.".

It also depends upon the purposes of experiment, usually, a difference is marked "significant", when the gap between the two sample means points to or signifies a real difference between parameters of the population from which the sample are drawn.

The research workers as the experimenters have an option to choose several arbitrary standards called levels of significance of which the .05 and .01 levels are most often used. The confidence with which an experimenter or research worker, rejects or retains (accept) a null hypothesis, depends upon the level of significance adopted.

You carefully look at the table Z distribution presented in unit 1, table no. 1.6.1, you will find that at the point ±1.96 the total 95% cases fall.

If we take ±1.96 at the base line of normal distribution curve as two points, we find that total 95% area of the curve lie between these two points.

Remaining 5% area lies left on to the right side of the curve i.e. 2½ % area lies to the left and 2 ½ % area lies to the right side of the curve.

If the Z Score of the mean difference which is also known as "C. R. value of t ratio" is 1.96 or below, it means the difference of the two means lies within the acceptance area of the normal distribution of the sampling distribution of the area differences.

Hence the null hypotheses (Ho) are retained. "CR" or "t ratio" is higher than 1.96, means the mean difference falls within the area of rejection, hence the null hypotheses (Ho) is rejected.

Further, you see the table 1.6.1 again, you will find that at the point ±2.58ó on the base of the normal distribution curve total 99% area of the curve or cases lie within the range -2.58 ó to + 2.58ó.

Only 1% area of the curve or cases lie beyond these two units. If the CR value or "t ratio" is below the value of 2.58, i.e. within the area of acceptance of 99% level the obtained mean difference is significant at .01 level or 99% level.

If the CR value of t. ratio is obtained above to the 2.58% , the null hypothesis (Ho) said to be rejected at 99% level or .01 level of significance.

2.9.3 The Null Hypothesis

In the above paragraphs a term Null Hypothesis is used in relation to determining of the significance of difference between the two means. Before we proceed further, it is essential to know about the null hypothesis and its rate in determining the significance of the difference of "Zero difference" or "No difference" in the relative specific parameters of the population and symbolically it is denoted as Ho.

Hypothesis is a suggested or pre determined relation of a problem which is tested on the basis of the evidences collected. Null hypothesis is a useful tool in testing the significance of the difference.

The null hypothesis a states that there is no true difference between two population means, and that the difference found between sample means, therefore, are only by chance i.e. accidental and unimportant.

The null hypothesis is based on the simple logic that a man is innocent until he is proved guilty. It constitutes or brings a challenge before the experimenter to call the necessary evidences to reject or retain the null hypothesis which he has framed.

After rejection of the null hypothesis automatically the alternative hypotheses will he accepted, For Example: In a study of Linguistic ability of boys and girls of group 14+years, the researcher has framed the following two hypothesis-

Ho: There is no difference in the means of linguistic ability scores of male and female adolescents of age group 14+ years.

$H_{A:}$ The mean of the linguistic ability scores is in favour of the adolescents girls than the boys of age group 14+ -16 + years.

It is obvious that if the null hypothesis (Ho) is rejected on the basis of statistical treatment made on the related evidences collected, the alternative hypothesis (H_A) will be accepted. If the null hypothesis (Ho) is accepted on the basis of the evidences collected, in such condition the alternative hypothesis (H_A) will be rejected.

2.9.4 Basic Assumption of Testing of Significance difference between the Two Sample Means

The formula which is used to test the significance of the difference between the two sample means (see 2.9.1) is based on certain basic assumptions. The assumptions are as under:

1) The variable or the trait being measured or studied is normally distributed in the universe.

2) There is no difference in the Means of the two or more populations. i.e. My=Mz If there is a violation or deviation in the above assumptions in testing the significant difference in the two means, we can not use "C.R" or "t" test of significance. In such condition, there are other methods, which are used for the purpose.

3) The samples are drawn from the population using random method of sample selection.

4) The size of the sample drawn from the population is relatively large.

2.9.5 Two Tailed and One Tailed Test of Significance

Under the null hypothesis, difference between the obtained sample means from two populations i.e. M_1 and M_2 may be either plus or minus and is often in one direction.

In yet another case, the differences between true parameters may be difference of Zero i.e. M_1-M_2=0, or Mz_{DM}=0, so that in determining probabilities we consider both traits of the sampling distribution. This two tailed test, as it is sometimes called, is generally used when we wish to discover whether two groups have conceivably been drawn from the same population with respect to the trait being measured.

In many research studies or experiments the primary concern is with the direction of the difference rather than with its existence in absolute terms. This situation arises when we are not interested in negative differences or in the losses made as this has no importance practically. However, we are much interested in the positive directions i.e. gains or growth or developments. For example, suppose we want to study the effect of extrinsic motivation on solving the mathematical problems or in sentence construction, it is unlikely that extrinsic motivation leads to loss in either solving the mathematical problems correctly or framing the sentences correctly.

Thus here only we are interested the positive effect of motivation and we study both gain made by the learners in solving the mathematical problems or constructing the sentences correctly. In such condition we use one tail of the normal probability curve that is the positive side and the Z or ó values will be changed after 95% and 99% level of significance far. In case of large samples the Z or ó values for 95% level it becomes 12.33 ó in place of 2.586.

2.9.6 Uncorrelated (Independent) and Correlated (Dependant) Sample Means

When we are interested to test whether two groups differ significantly on a trait or characteristics measured, the two situations arises with respect to differences between means

1) Uncorrelated or Independent two sample means
2) Correlated or Dependent two sample means

The two sample means are uncorrelated or independent when computed from different samples selected by using random method of sample selection from one population or from different populations or from uncorrelated tests administered to the same sample.

The two sample means are correlated when a single group of population is tested in two different situations by using the same test. In other words when one test is used on a single group before the experiment and after the experiment or when the units of the Group or the population from which two sample are drawn are not mutually exclusive.

In the latter situation, the modified formula for calculating the standard error of the difference of two sample means is applied.

Thus to test the significance of sample means, there are always following four s situations:

1) Two Large Independent Sample. i.e. when N_1 and N_2 >30

2) Two Small Independent Samples i.e. when N_1 and N_2 <30

3) Two Large correlated samples.

4) Two Small correlated samples.

2.10 SIGNIFICANCE OF THE TWO LARGE INDEPENDENT OR UNCORRELATED SAMPLE MEANS

The formula for testing the significance of two large independent sample means is as follows:

$CR = M_1 \sim M_2$ where $\sigma_{DM} = \sqrt{\sigma M_1^{2} + \sigma M_2^{2}}$ or

$= \sqrt{\sigma 1^2} \ / \ N1 + \sigma 2^{2} / \ N2$

Self Assessment Questions

1) What do you mean by significance of the difference in two means?

....

....

....

....

2) Define Standard Error of the difference of the two sample means.

....

....

....

....

3) Define Sampling distribution of the differences of Means of two Samples.

....

....

....

....

4) What should be the mean value of sampling distribution of the difference of the means?

....

....

....

....

5) What indicates S.E.$_{DM}$ or ____$_{DM}$?

....

....

....

....

6) What do you mean by Ho, Define it.

....

....

....

....

7) What are the assumptions on which testing of the difference of two Mean is based?

....

....

....

....

8) What do you mean by One Tail Test and Two Tail Test? When these two tails are used?

....

....

....

....

9) What is meant by uncorrelated and correlated sample means?

....

....

....

....

Thus

CR = $M_1 \sim M_2$ / ó D

$$= \frac{M_1 \sim M_2}{\sqrt{\sigma 1^2 + \sigma 2^2}}$$

N_1 N_2

Where

CR : Critical Ratio

M_1 : Mean of the Sample or Group 1

M_2 : Mean of the Sample or Group 2

ó¹ : Standard Deviation of the Scores of sample 1

ó² : Standard Deviation of the Scores of sample 2

N_1 : Number of cases in Sample 1

N_2 : Number of cases in sample 2

Example 8: An Intelligence test was administered on the two groups of Boys and Girls. These two groups were drawn from the two populations independently by using random method of sample selection. After administration of the test, the following statistics was obtained

Groups	N	M	Σ
Boys	65	52	13
Girls	60	48	12

Determine the difference between the mean values of Boys and Girls significant?

Solution

In the given problem, the two samples are quite large and independent. Therefore, to test the significance difference in the mean values of Boys and Girls. First we have to determine the null hypothesis which is

$Ho=M_B=M_G$ i.e.

There is no significant difference in the mean value of the Boys and Girls and the two groups are taken from the same population

$$\text{C.R.} = \frac{M_1 \sim M_2 - 0}{\sigma D}$$

$$= \frac{(M_1 \sim M_2)}{\frac{\sqrt{\sigma 1^2}}{N_1} + \frac{\sqrt{\sigma 2^2}}{N_2}}$$

$$\text{C.R.} = \frac{(52-48)}{\frac{\sqrt{13^2}}{65} + \frac{\sqrt{12^2}}{60}} = \frac{4}{\frac{\sqrt{169}}{65} + \frac{\sqrt{144}}{60}} = \frac{4}{\sqrt{5}}$$

Or C.R. = 1.79

df $= (N_1\text{-}1) + (N_2\text{-}1)$

$= (65\text{-}1) + (60\text{-}1)$

$= 123$

To test the null hypothesis, which is framed, we will use two tail test. In the "t" distribution table (sub heading no. 2.5.2) at 123 df the "t" value at .05 level and .01 level is 1.98 and 2.62 respectively (The "t" table has 100 and 125 df, but df 123 is not given, therefore nearest of 123 i.e. 125df is considered). The obtained t value (1.79) is much less than these two values, hence it is not significant and null hypothesis is accepted at any level of significance.

Interpretation of the Results

Since our null hypothesis is retained, we can say that Boys and Girls do not differ significantly in their level of intelligence. Whatever difference is observed in the obtained mean values of two samples is due to chance factors and sampling fluctuations. Thus we can say with 99% level of confidence that no sex difference exists in the intelligence level of the population.

2.11 SIGNIFICANCE OF THE TWO SMALL INDEPENDENT ON UNCORRELATED SAMPLE MEANS

When the N's of two independent samples are small (less than 30), the S.E. $_{DM}$ or -ó$_{DM}$ (standard error of the difference between two means) should depend upon the Standard Deviation (S.D. or ó) values computed by using the correction

The formula used to test the significance of the difference of two means of small independent samples is :

$$t = \frac{M_1 \sim M_2}{S.E_{DM}}$$

Where

$$S.E_{DM} = S.D \frac{\sqrt{N_1 + N_2}}{N_1 \times N_2}$$

$$\text{And } S.D = \frac{\sqrt{\Sigma(x1\text{-}M1)^2 + \Sigma(x2\text{-}M2)^2}}{(N_1 - 1)(N_2 - 1)}$$

For simplification the above formula can also be written as

$$t = \frac{M1 \sim M2}{\frac{\Sigma d2 + \Sigma d2^2}{N1+N2-2} \times \frac{N1+N2}{N1N2}} \quad \text{— (i)}$$

Where

$D_1 = (S_1\text{-}M_1)$, and

$D_2 = (x_2\text{-}M_2)$

Here, X_1 and X_2 are the new scores of two groups, M_1 and M_2 are given in relation to the two samples or groups having the small number units or cases.

When the raw data are not given and we have statistics or the estimates of two small size sample, in such condition, we use the formula-

The corresponding Mean values of the scores of the two groups N_1 and N_2 are the number of the units or cases in the two groups t is also a critical ratio in which more exact estimate of the $ó_{DM}$ is used. Here 't' in place C.R. is used because sampling distribution of "t" is not normal when N is small i.e. <30, "t" is a critical ratio (C.R.), but all C.R's are not "t's.

$$t = \frac{M_1 \sim M_2}{\sqrt{\sigma 1^2 \left(N_1 - 1\right) \sigma 2^2 \left(N_2 - 1\right) \times \left(N_1 + N_2\right) / N_1 N_2}} \qquad \text{.......(ii)}$$

Where

M_1 = Mean of the scores of sample -1

M_2 = Mean of the scores of sample -2

$ó^1$ = Standard Deviation of the scores of sample-1

$ó^2$ = Standard Deviation of the scores of sample-2

N_1 = Number of units or cases on the sample-1

N_2 = Number of units or cases in the sample-2

For more clarification study the following examples very carefully.

Example 9: An attitude test regarding a vocational course was given to 10 urban boys and 5 rural boys. The obtained scores are as under-

Urban Boys (x1) = 6, 7, 8, 10, 15, 16, 9, 10, 0, 9

Rural Boys (x2) = 4, 3, 2, 1, 5

Determine at .05 level of significance that its there a significant difference in the attitude of boys belonging to rural and urban areas in relation to a vocational course?

Solution

Ho=b1=b2 : H1=b1 $\neq$ b2

Level of significance = .05

For acceptance or rejection of null hypothesis at .05 level of significance, the two tail test is used.

Thus

Urban Boys			**Rural Boys**		
X1	d1=(x1-m1)	d12	X2	d2=(X2-M2)	d22
6	-4	16	4	+1	1
7	-3	9	3	0	0
8	-2	4	2	-1	1
10	0	0	1	-2	4
15	+5	25	5	+2	4
16	+6	36			
9	-1	1	$\sum x_2=15$		
10	0	0			$\sum d_2^2=10$
10	0	0	M= 15/5		
9	-1	1	M= 3		

$\sum x_1=100$ $\qquad \sum d1^2=92$

M= $\sum$x/N

= 100/10

M=10

We know that

$$t = \frac{M_1 \sim M_2}{\sqrt{\frac{\Sigma d1^2 + \Sigma d2^2}{N_1+N_2-2} \times \frac{N_1+N_2}{N_1 N_2}}}$$

$$= \frac{10-3}{\sqrt{\frac{92+10}{10+5-2} \times \frac{10+5}{10\times 5}}}$$

$$= \frac{7}{\sqrt{7.8 \times 0.30}} = \frac{7}{\sqrt{2.34}} = \frac{7}{1.54}$$

Or t = 4.46

df = $(N_1\text{-}1)+(N_2\text{-}1)$

= 9+4

=13

In "t" distribution table (table 2.5.1), the t value for 13 df at .05 level is 2.16. The obtained t value 4.46 is much greater than this value. Hence null hypothesis is rejected.

Interpretation of the Result

Our null hypothesis is rejected at .05 level of significance for 13 df. Thus we can say that in 95% cases significant difference in the attitude of the urban and rural boys regarding a vocational course. There are only 5% chances out of 100 that the two groups have same attitude towards a vocational course.

Example 10: music interest test was administered on 15 + years did boys and girls sample taken independently from the two populations. The following statistics was obtained:

	Mean	S.D.	N
Girls	40.39	8.69	30
Boys	35.81	8.33	25

Is the mean difference is in favour of girls?

Solution:

$Ho = b_1 = b_2$

$H_1 = b_1 \neq b_2$

In the given problem, the row scores of the two groups are not given. Therefore we will use the following formula for testing of the difference of means of two uncorrelated sample means:

$$t = \frac{M_1 \sim M_2}{\sqrt{\frac{\Sigma 1^2 (N_1 - 1) + \Sigma 2^2 (N_2 - 1)}{N_1 + N_2 - 2} \times \frac{N_1 + N_2}{N_1 \times N_2}}}$$

$$t = \frac{40.39 - 35.81}{\sqrt{\frac{(8.69)^2 (30\text{-}1) + (8.33)^2 (25\text{-}1)}{30+25-2} \times \frac{30+25}{30 \times 25}}}$$

$$= \frac{4.58}{\sqrt{75.516 \times 29 + 69.389 \times 24 \times 55}}$$

$$= \frac{4.58}{\sqrt{7274 \times .073}} = \frac{4.58}{2.309}$$

Or t $= 1.98$

d f $= (N_1 - 1) + (N_2 - 1) = 53$

In the t distribution table for 53 df the t value at .05 level is 2.01. Our calculated t value 1.98 is less than this value. Therefore, the null hypothesis is retained.

Interpretation of the Results

Since our null hypothesis is accepted at .05 level of significance. Therefore it can be said that in 95 cases out of 100, there is no significant difference in the mean values of boys and girls regarding their interest in music. There are only 5% chances that the two groups do not have equal interest in music. Hence with 95% confidence, we can say that both boys and girls have equal interest in music. Whatever difference is deserved in the mean values of the groups is by chance or due to sampling of fluctuations.

2.12 SIGNIFICANCE OF THE TWO LARGE CORRELATED SAMPLES

In some of the experimental studies a single group is tested in two different conditions and the observations are in pairs. Or two groups are used in the experimental condition, but they are matched by using pairs method. In these conditions, a modified formula of standard error of the difference of means is used. Therefore the formula for testing of the difference of two means of large correlated samples is –

$$t = \frac{M_1 \sim M_2}{\sqrt{\sigma M_1^{\ 2} + \sigma M_2^{\ 2} - 2r_{12}\sigma M_1 \ \sigma M_2}}$$

In the formula

M_1 = Mean of the scores of sample -1

M_1 = Mean of the scores of sample -2

óM_1 = Standard Error of the Mean of sample-1

i.e. $óM_1 = \frac{\Sigma 1}{\sqrt{N_1}}$

$óM_2 = \frac{\Sigma 2}{\sqrt{N_1}}$

and r1,2 = correlation between two sets of scores.

For more classification go through the following examples

Example 11: An Intelligence test was administered on a group of 400 students twice after an interval of 2 months. The data obtained are as under-

		M	S.D
Testing –I	:	25	8
Testing-II	:	30	5
N	:	400	
r_{12}	:	65	

Test if there is a significant difference in the means of intelligence scores obtained in two testing conditions.

Solution:

Ho => b1=b2 and H1=>b1 $\neq$ b2

$$\therefore \ t = \frac{M_1 \sim M_2}{\sqrt{\sigma M_1^{\ 2} + \sigma M_2^{\ 2} - 2r_{12}\sigma M_1 \ \sigma M_2}}$$

According the formula all values are given except S.E of means ($\sum$ M). Therefore first we have to calculate standard errors of the means of the two sets of scores

$$\therefore \ óM_1 = \frac{\Sigma 1^2}{\sqrt{N_1}} = \frac{8^2}{\sqrt{400}} = \frac{64}{20}$$

Or $\sigma M_1 = 3.20$

Similarly

$$\sigma M_2 = \frac{\sigma_2{}^2}{\sqrt{N_2}} = \frac{5^2}{\sqrt{400}} = \frac{25}{20}$$

Or $\sigma M_2 = 1.25$

Thus

$$t = \frac{30-25}{\sqrt{(3.20)^2 + (1.25)^2 - 2 \times .65 \times 3.20 \times 1.25}}$$

$$= \frac{5}{\sqrt{10.24 + 1.5625 - 5.20}}$$

$$= \frac{5}{\sqrt{6.6025}} = \frac{5}{2.57}$$

$t = 1.95$

df = N-1 = 400-1

df = N-1= 400-1 (In the example N is same i.e. the single group is tested in two different time intervals)

a df = 399

According to "t" distribution table (Table no-2.5.1) the value of t for 399 df at .01 level is 2.59. Our calculated value of t is 1.95, which is smaller than the value of t given in "t" distribution table. Hence the obtained t value is not significant even at .05 level. Therefore our null hypothesis is retained at .01 level of significance.

Interpretation of the Results

Since the obtained t value is found insignificant level for 399 df; thus the difference in the mean values of the intelligence scores of a group, tested after an interval of two months is not significant in 99 conditions out of 100, there is only 1% hence that the difference in two means is significant at .01 level.

Example 12: In a vocational training course an achievement test was administered on 64 students at the time of admission. After training of one year the same achievement test was administered. The results obtained are as under:

		M	σ
Before Training	:	52.50	7.25
After Training	:	58.70	5.30

Is the gain, after training significant?

Solution:

$H_o = b_1 = b_2$ (The gain after training is insignificant)

$H_1 = b_1 \neq b_2$

(*Note:* Read the problem carefully, here we will use one tail test rather to use two tail test. Because here we are interested in gain due to training, not in the loss. That is we are interested in one side of the B.P.C which is +ve side. 99% confidence and .05 for 95% confidence. See the table no-2.5.1 carefully and read the footnote)

We know that formula of testing the difference between two large correlated means is–

$$\text{Formula } t = \frac{M_1 \sim M_2}{\sqrt{\sigma M_1^{\,2} + \sigma M_2^{\,2} - 2r_{12}\sigma M_1 \; \sigma M_2}}$$

Where

$$\acute{o}M_1 = \frac{\sigma 1}{\sqrt{N}} = \frac{7.25}{\sqrt{100}} = \frac{7.25}{10}$$

Or $\acute{o}M_1 = .725$

$$\text{And } \acute{o}M_2 \quad \frac{\sigma 2}{\sqrt{N}} = \frac{530}{\sqrt{N}} = \frac{5.30}{\sqrt{100}} = \frac{5.30}{10} = .53$$

Or $\acute{o}M_2 = .53$

$$t = \frac{58.70\text{–}52.50}{\sqrt{(.725)^2 + (.53)^2 - 2 \times .50 \times .725 \times .53}}$$

$$= \frac{6.2}{\sqrt{0.4223}} = \frac{6.2}{.65}$$

$t = 9.54$

df = (100-1)

= 99

In the‘t’ distribution table (table No. 2.5.1) at .02 level the t value for 99 df is 2.36 and out obtained t value is 9.54, which is much greater than the “t” value of the table. Thus the obtained t value is significant at 99% level of significance. Therefore our null hypothesis is rejected.

Interpretation of the Results

Since the obtained “t” value is found significant at .02 level for 99df. Thus we can say that gain on the achievement test made by the students after training is highly significant. Therefore we can say with 99% confidence that given vocational training is quite effective. There is only 1 chance out of hundred, the vocational training is ineffective.

2.13 SIGNIFICANCE OF TWO SMALL CORRELATED MEANS

In case of determining the significance difference between the two correlated small sample mean we have two methods, which are as under

i) **Direct Method:** i.e. we have to calculate Mean Values standard deviation values of the two groups and the coefficient of correlation (r_{12}) between the scores of two groups. In such condition the formula used to test significant difference in the two means is –

$$t = \frac{M_1 \sim M_2}{\frac{\sqrt{\sigma 1^2}}{N-1} + \frac{\sqrt{\sigma 2^2}}{N-1} - 2r_{12}\frac{\sigma 1 \sigma 2}{N-1}}$$

$$t = \frac{M_1 \sim M_2}{\sqrt{Sm_1^{\ 2} + Sm_2^{\ 2} - 2r_{12}Sm_1Sm_2}}$$

Where

$$Sm_1 = \frac{\sigma 1}{\sqrt{N-1}} \text{ (standard error of the small sample mean)}$$

$$Sm_2 = \frac{\sigma 2}{\sqrt{N-2}}$$

ii) **Difference Method:** In this method we have the raw data of two small groups or sample and we are not calculate coefficient of correlation (r_{12}) between the two sets of scores.

Examples 13: A pre test and past test are given to 12 subjects. The scores obtained are as under–

S. No.-	1	2	3	4	5	6	7	8	9	10	11	12
Pre-Test:	42	50	51	26	35	42	60	41	70	38	62	55
Past-Test:	40	62	61	35	30	52	68	51	84	50	72	63

Determine if the gain on past test score significant?

Solution:

S.No. of Subjects	Post Test X_1	Pre Test X_2	Difference $(X_2\text{-}X_1)$	D-MD d	d^2
1	40	42	-2	-10	100
2	62	50	12	4	16
3	61	51	10	2	4
4	35	26	9	1	1
5	30	35	-5	-13	169
6	52	42	10	2	4
7	68	60	8	0	0
8	51	41	10	2	4
9	24	70	14	6	36
10	50	38	12	4	16
11	72	62	10	2	4
12	63	55	8	0	0
			$\sum D = 96$		$\sum d^2 = 354$

$$MD = \frac{\sum D}{N}$$

$$SD = \frac{\sqrt{\sum d^2}}{N-1}$$

$$= \frac{\sqrt{354}}{14}$$

Or SD= 5.67

$$\therefore \quad SE_{DM} = \frac{\sigma D}{\sqrt{N}}$$

Where

SE_{DM} = Standard Error of the Mean of Differences.

ΣD = Standard Deviation of the Differences

And N = Total No. of cases.

$$\text{Thus } SE_{DM} = \frac{5.67}{\sqrt{12}} = \frac{5.67}{3.464}$$

$$= 1.631$$

$$\therefore \; t = \frac{MD}{SE_{DM}} = \frac{8}{1.637} = 4.88, \; df = 11$$

In the "t" distribution table (Table 2.5.1 subheading 2.5.2) for 11 df at .02 level the value is 2.72 and our calculated value of t (4.88) is much greater than the table value. Therefore the null hypothesis is rejected at .01 level of significance.

Interpretation of the Results

Since our null hypothesis is rejected at.01 level of significance, therefore we can say that the gain made by the subject on past test is real in 99 cases out of 100. There are only 1% chance that the gain shown by the subjects in cue to chance factors as by sampling fluctuations.

2.14 POINTS TO BE REMEMBERED WHILE TESTING THE SIGNIFICANCE IN TWO MEANS

When you compare the means of two groups or to compare the means of a single group tested in two different situations or conditions and to know, whether the difference found in the two means is real and significant, or it is due to chance, factors, you should keep in mind the following steps as a process of testing the difference between two means.

i) Set up null hypothesis (Ho) and the alternative hypothesis (H1), according to the requirements of the problem.

ii) Decide about the level of significance for the test, usually in behavioural or social science, .05 and .01 levels are taken into consideration for acceptance or rejection of the null hypothesis.

iii) Decide whether one tailed or two tailed test of significance for independent or the correlated means.

iv) Decide whether the large or samples are involved in the problem or in the experiment.

v) Calculate either C.R value or "t" ratio value as per nature and size of the samples, by using the formulas discussed on the precious pages.

vi) Calculate degree of freedom (df). It should be N1+N2-2 for independent t or uncorrelated samples. While in case of correlated samples it should be N-1.

vii) Consult the "t" distribution table with df keeping in mind the level of significance, which we have decided at step number-11.

viii) Compare the calculated value of "t" with the "t" value given in the table with respect to df and level of significance.

ix) Interpret the Results:

If null hypothesis (Ho) is rejected, there is a significant difference between the two means.

If null hypothesis is accepted, there is no significant difference in the two means. Whatever the difference exists it has arisen due to sampling fluctuations or chance factors only.

Self Assessment Questions

1) How you can define "Critical Ratio" and "t" Ratio?

....

....

....

....

2) What is the difference between "CR" and "t"?

....

....

....

....

3) What is the difference between C.R. and Z Score?

....

....

....

....

4) How can you define the standard Error of the Difference of Means?

....

....

....

....

5) What formula you will use in the following conditions:

a) When two independent large samples are given.

....

....

....

b) When two correlated large samples are given.

....

....

....

c) When two independent small samples are given.

....

....

....

d) When two small independent samples are given.

....

....

....

6) What do you mean by independent samples?

....

....

....

....

7) What do you mean by correlated samples?

....

....

....

....

2.15 ERRORS IN THE INTERPRETATION OF THE RESULTS, WHILE TESTING THE SIGNIFICANT DIFFERENCE BETWEEN TWO MEANS

While interpreting the results obtained from the test of significance of a single mean or the difference between two means, we should take case and not to depend much on the statistical results obtained. Generally while interpreting the results, we may make following two type of Error.

Type-I Error or á Errors:

This Error occurs when the null hypothesis is true, but we reject the same by marking significant difference between the two means.

Type-II Error or â Error:

This error occurs when the null hypothesis is wrong or to be rejected while the same is retained.

This probability of occurrence of type-II error or â error due to finding high level of significance i.e. above to the .01 level of significance which may be .001 or the above.

2.16 LET US SUM UP

The Standard error of the estimates or statistics or sample statistical measures ($S.E._M$) consists –

Error of Sampling, and

Error of Measurement

Fluctuations from sample to sample, the so called sampling error or errors of sampling are not to be thought of as mistakes, features, and the like, but as variations arising from the fact that no two samples are over exactly alike.

Mean values and standard deviations (Σ's) obtained from random samples taken from a population are estimates of their parameters (the true statistical measurements of the population) and the standard error ($S.E_M$) measures the tepidness of these estimates.

If the $S.E_M$ is large, it does not follow necessarily that the obtained statistics is effected by a large sampling error, as much of the variations may be due to error of measurements, when error of measurements are low i.e. the measuring tools or tests are highly reliable, a large $S.E_M$ indicates considerable sampling error.

In the comparative studies or the experiments it is to decide whether the obtained differences of such magnitude is attributed to chance factor or sampling variations or it really exists? For such decisions the standard error of the difference is considered.

The critical or "t" ratio are nothing, but these are the Z scores , which tells how far the sample mean difference derivates to the population mean difference on a normal distribution curve.

"C.R" and "t" are the ratio of the mean difference of the two groups and the standard error of the mean differences.

While deciding the significance of any statistical measure or the difference on the means of two samples or two populations, the degree of freedom and levels of confidence are considered, and in the light of these two we either accept the null hypothesis or to reject the same.

While taking the decision a care is to be taken, so that type-I and type-II error should not be occur.

2.17 UNIT END QUESTIONS

1) Explain the term "Statistical Inference". How is the statistical inference is based upon the estimation of parameters.

2) Indicate the role of standard error for statistical generalisation.

3) Differentiate between significance of statistics and confidence interval of fiduciary limits.

4) Enumerate the various uses of Standard Error of the statistics.

5) What type of errors can occur while interpreting the results based on test of significance? How we can overcome these errors?

6) A Sample of 100 students with mean score 26.40 and SD 5.20 selected randomly from a population. Determine the .95 and .99 for confidence intervals for population true mean.

7) A small sample of 10 cases with mean score 175-50 and $\sum$=5.82 selected randomly. Compute finding limits of parameter mean at .05 and .01 level of confidence.

8) The mean and standard deviation of the intelligence scores obtained on a group of 200 randomly selected students are 102 and 10.20 respectively. How dependable is mean I.Q. of the students?

The following are the data for two independent samples :

	N	M	S.D.
Boys	60	48.50	10.70
Girls	70	53.60	15.40

Is the difference in the mean values of Boys and Girls significant.

A reasoning ability test was given to 8 urban and 6 rural girls of the same Class. The data obtained are differ significantly in there reasoning ability.

Groups	Scores
Urban Girls	16,9,4,23,19,10,5,2
Rural Girls	20,5,1,16,2,4.

The observations given below obtained on 10 subjects in a experiment of Pre and Post test. Is gain trade by the students on post test significant?

Subjects	1	2	3	4	5	6	7	8	9	10
Scores on Pre Test	5	15	9	11	4	9	8	13	6	16
Scores on Post Test	7	9	4	15	6	13	9	5	6	12

9) A group of 10 students was given 5 trials on a test of physical efficiency. Their score on the I and V trials are given below. Test whether there is a significant effect of practice on the improvement made in first to fifth trial.

Subject	A	B	C	D	E	F	G	H	I	J
Trial I	15	16	17	20	25	30	17	18	10	12
Trial V	20	22	22	25	35	30	21	23	17	20

10) A group of 35 students randomly selected was tested before and after experimental treatment. The observations obtained are as under:

	M	ó
Pre Test	15.5	5.2
Post Test	21.6	4.8

Coefficient of Correlation between The scores of Pre and Post Test	0.70

Find out the groups is different significantly on the two testing conditions.

2.18 POINTS FOR DISCUSSION

1) What will happen on the standard error of sample mean if

   a) Sample is homogeneous and large

   b) Sample is heterogeneous and large

   c) Sample is heterogeneous and small

   d) Sample is homogeneous as well as small

2) Differentiate between "t" distribution and Z distribution. What is the basic difference between "t" Test and Z Test.

3) When the "t" distribution and "Z" distribution become conside

4) The necessity of a theoretical distribution model for estimation.

2.19 SUGGESTED READINGS

Aggarwal Y.P. (1990) *Statistical Methods – Concepts Application and Computation*. Sterling Publications Pvt Ltd. New Delhi.

Walker . H.M. and Lev. J. (1965). *Statistical Inference.* Oxford and I B H Publishing Co. Calcutta.

UNIT 3 ONE WAY ANALYSIS OF VARIANCE

Structure

3.0 INTRODUCTION

In the foregoing unit you have learned about how to test the significance of a mean obtained on the basis of observations taken from a group of persons and the test of significance of the differences between the two means. No doubt the test of significance of the difference between the two means is a very important technique of inferential statistics, which is used to test the null hypothesis scientifically and help to draw concrete conclusion. But its scope is very limited. It is only applicable to the two sets of scores or the scores obtained from two samples taken from a single population or from two different populations.

Now imagine if we have to compare the means of more than two populations or the number of groups, then what would happen? Can we apply successfully the Critical Ratio Test (CR) or the t test? The answer is yes, but not convenient to apply CR test or t test. The reason can be stated with an example. Suppose we have three groups A,B & C and we want to compare the significance difference in the means of the three groups, then first we have to make the pairs of groups e.g. A and B, then B and C, and then A and C and apply C.R. test or t test as the conditions required. In such condition we are to calculate three C.R. values or t values instead of one.

Now suppose we have eight groups and want to compare the difference in the means of the groups, in such condition we have to calculate 28 C.R. or t values as the condition may require.

It means when there are more than two groups say 3, 4, 5 ….. and k, it is not easy to apply 'C.R.' or 't' test of significance very conveniently.

Further 'C.R.' or 't' test of significance simply consider the means of two groups and test the significance of difference exists between the two means. It has no concern in the variance that exist in the scores of the two groups or variance of the scores from the mean value of the groups.

For example let us say that A reaction time test was given to 5 boys and 5 girls of age group 15+ yrs. The scores were obtained in milliseconds are as given in the table below.

Groups	Reaction time in M. Sec					Sum	Mean
Girls	15	20	5	10	35	85	17M.Sec.
Boys	20	15	20	20	10	85	17M.Sec.

From the mean values shown in the table we can say that the two groups are equal in their reaction time and the average reaction time is 17 M. Sec. In this example, if we apply 't' test of significance, we will find, the difference in the two means insignificant and our null hypothesis is retained.

But if we look carefully to the individual scores of the reaction time of boys and girls, we will find that there is a difference in the two groups. The group of girls is very heterogeneous in their reaction time in comparison to the boys.

As the variation between the scores is ranging from 5 to 30 and deviation of scores from mean varies from 12 M. Sec. to 18 M. Sec.

While the group of boys is more homogeneous in their reaction time, as the variation in the individual scores is ranging from 5 to 10 and deviation of the scores from mean is 3 M. Sec to 7 M. Sec therefore group B is much better in their reaction time in comparison to the group A.

From, this example, you have seen that the test of significance of difference between the two means, some time lead us to draw wrong conclusion and we may wrongly retain the null hypotheses, though it should be rejected in real conditions.

Therefore, when we have more than two, say three or four or so forth and so on, the 'CR' or 't' test of significance are not very useful. In such condition, 'F' test is more suitable and it is known as one way analysis of variance. Because we are testing the significance difference in the average variance exists between the two or more than two groups, instead to test the significance of the difference of the means of the groups.

In this unit we will be dealing with F test or the analysis of variance.

3.1 OBJECTIVES

After going through this unit, you will be able to:

- Define variance;
- Differentiate between variance and standard deviation;
- Define analysis of variance;
- Explain when to use the analysis of variance;

- Describe the process of analysis of variance;
- Apply analysis of variance to obtain 'F' Ratio and to solve related problems;
- Analyse inferences after having the value of 'F' Ratio;
- Elucidate the assumptions of analysis of variance;
- List out the precautions while using analysis of variance; and
- consult the 'F' table correctly and interpret the results.

3.2 ANALYSIS OF VARIANCE

The analysis of variance is an important method for testing the variation observed in experimental situation into different part each part assignable to a known source, cause or factor.

In its simplest form, the analysis of variance is used to test the significance of the differences between the means of a number of different populations. The problem of testing the significance of the differences between the number of means results from experiments designed to study the variation in a dependent variable with variation in independent variable.

Thus the analysis of variance, as the name indicates, deals with variance rather than with standard deviations and standard errors. It is a method of dividing the variation observed in experimental data into different parts, each part assignable to a known source, cause or factor therefore

$$F = \frac{\text{Variance between the groups}}{\text{Variance within the groups}} = \frac{\sigma^2\text{Between the groups}}{\sigma^2\text{Within the groups}}$$

In which σ^2 is the population variance.

The technique of analysis of variance was first devised by Sir Ronald Fisher, an English statistician who is also known as the father of modern statistics as applied to social and behavioural sciences. It was first reported in 1923 and its early applications were in the field of agriculture. Since then it has found wide application in many areas of experimentation.

3.2.1 Meaning of the Variance

Before to go further the procedure and use of analysis of variance to test the significance difference between the means of various populations or groups at a time, it is very essential, first to have the clear concept of the term variance.

In the terminology of statistics the distance of scores from a central point i.e. Mean is called deviation and the index of variability is known as the mean deviations or standard deviation (σ).

In the study of sampling theory, some of the results may be some what more simply interpreted if the variance of a sample is defined as the sum of the squares of the deviation divided by its degree of freedom (N-1) rather than as the mean of the squares deviations.

The variance is the most important measure of variability of a group. It is simply the square of S.D. of the group, but its nature is quite different from standard deviation, though formula for computing variance is same as standard deviation (S.D.)

$$\therefore \text{ Variance} = \text{S.D.}^2 \text{ or } \sigma^2 = \frac{\Sigma(X-M)^2}{N}$$

Where X : are the raw scores of a group, and

M : Mean of the raw scores.

Thus we can define variance as **"the average of sum of squares of deviation from the mean of the scores of a distribution."**

3.2.2 Characteristics of Variance

The following are the main features of variance:

- The variance is the measure of variability, which indicates the among groups or between groups difference as well as within group difference.
- The variance is always in plus sign.
- The variance is like an area. While S.D. has direction like length and breadth has the direction.
- The scores on normal curve are shown in terms of units, but variance is a area, therefore either it should be in left side or right side of the normal curve.
- The variance remain the some by adding or subtracting a constant in a set of data.

Self Assessment Questions

1) Define the term variance.

....

....

....

....

2) Enumerate the characteristics of Variance.

....

....

....

....

3) Differentiate between standard deviation and variance.

....

....

....

....

4) What do you mean by Analysis of Variance? Why it is preferred in comparison to 't' test while determining the significance difference in the means.

....

....

....

....

3.2.3 The Procedure of Analysis of Variance (ANOVA)

In its simplest form the analysis of variance can be used when two or more than two groups are compared on the basis of certain traits or characteristics or different treatments of simple independent variable is studied on a dependent variable and having two or more than two groups.

Before we discuss the procedure of analysis of variance, it is to be noted here that when we have taken a large group or a finite population, to represent its total units the symbol 'N' will be used.

When the large group is divided into two or more than two sub groups having equal number of units, the symbol 'n' will be used and for number of groups the symbol 'k' will be used.

Now, suppose in an experimental study, three randomly selected groups having equal number of units say 'N' have been assigned randomly, three kinds of reinforcement viz. verbal, kind and written were used. After a certain period, the achievement test was given to three groups and mean values of achievement scores were compared. The mean scores of three groups can then be compared by using ANOVA. Since there is only one factor i.e. type of reinforcement is involved, the situation warrants a single classification or one way ANOVA, and can be arranged as below:

Table 3.2.1

S.N.	**Group - A** Scores of Verbal Reinforcement		**Group - B** Scores of Kind Reinforcement		**Group – C** Scores of Written Reinforcement	
	X_a	X_a^2	X_b	X_b^2	Xc_1	X_{c1}^2
	Xa_1	(X_{a1}^2)	X_{b1}	(X_{b1}^2)	Xc_1	(X_{c1}^2)
	Xa_2	(X_{a2}^2)	X_{b2}	(X_{b2}^2)	Xc_2	(X_{c2}^2)
	Xa_3	(X_{a3}^2)	X_{b3}	(X_{b3}^2)	Xc_3	(X_{c3}^2)
	Xa_4	(X_{a4}^2)	X_{b4}	(X_{b4}^2)	Xc_4	(X_{c4}^2)
	Xa_5	(X_{a5}^2)	X_{b5}	(X_{b5}^2)	Xc_5	(X_{c5}^2)
	.	.	.	.	.	.
	.	.	.	.	.	.
	Xa_n	(X_{an}^2)	X_{bn}	(X_{bn}^2)	Xc_n	(X_{cn}^2)
Sum	$\sum Xa$	$\sum X_a^2$	$\sum X_b$	$\sum X_b^2$	$\sum Xc1$	$\sum X_{c1}^2$
Mean	$\frac{\sum Xa}{n}$ =Ma		$\frac{\sum Xb}{n}$ =Mb		$\frac{\sum Xc}{n}$=Mc	

To test the difference in the means i.e. MA, MB and MC, the one way analysis of variance is used. To apply one way analysis of variance, the following steps are to be followed:

Step 1 Correction tem Cx = $\frac{\left(\sum x\right)^2}{N} = \frac{\left(\sum x_a + \sum x_b + \sum x_x\right)^2}{n_1 + n_2 + n_3}$

Step 2 Sum of Squares of Total $SS_T = \sum x^2 - Cx$

$$= \sum x^2 - \frac{\left(\sum x\right)^2}{N}$$

$$= \left(\sum x_a^2 + \sum x_b^2 + \sum x_c^2\right) - \frac{\left(\sum x_3\right)}{N}$$

Step 3 Sum of Squares Among the Groups $SS_A = \frac{\left(\sum x\right)^2}{N} - Cx$

$$= \frac{\left(\sum x_a\right)^2}{n_1} + \frac{\left(\sum x_b\right)^2}{n_2} + \frac{\left(\sum x_c\right)^2}{n_3} - \frac{\left(\sum x\right)^2}{N}$$

Step 4 Sum of Squares Within the Groups $SS_W = SS_T - SS_A$

Step 5 Mean Scores of Squares Among the Groups $MSS_A = \frac{SS_A}{k-1}$

Where k = number of groups.

Step 6 Mean Sum of Squares Within the Groups $MSS_W = \frac{SS_W}{n-k}$

Where N = Total number of units.

Step 7 F Ratio i.e. F = $\frac{MSS_A}{MSS_W}$

Step 8 Summary of ANOVA

Table 3.2.2: Summary of ANOVA

Source of variance	Df	S.S.	M.S.S.	F Ratio
Among the Groups	k-1	SS_A	$\frac{SS_A}{K-1}$	$\frac{MSS_A}{MSS_W}$
Within the groups (Error Variance)	N-K	SS_W	$\frac{SS_w}{N-k}$	
Total	N-1			

The obtained F ratio in the summary table, furnishes a comprehensive or overall test of the significance of the difference among means of the groups. A significant F does not tell us which mean differ significantly from others.

If F-Ratio is not significant, the difference among means is insignificant. The existing or observed differences in the means is due to chance factors or some sampling fluctuations.

To decide whether obtain F-Ratio is significant or not we are taking the help of F table from a statistics book.

The obtained F-Ratio is compared with the F value given in the table keeping in mind two degrees of freedom k-1 which is also known as greater degree of freedom or df_1 and N-k, which is known as smaller degree of freedom or df_2. Thus, while testing the significance of the F ratio, two situations may arise.

The obtained F Ratio is Insignificant:

When the obtained F ratio is found less than the value of F ratio given in F table for corresponding lower degrees of freedom df_1 that is, k-1 and higher degree of freedom df that is, (df=N-K) (See F table in a Statistics Book) at .05 and .01 level of significance it is found to be significant or not significant. Thus the null hypothesis is rejected retained. There is no reason for further testing, as none of the mean difference will be significant.

When the obtained 'F Ratio' is found higher than the value of F ratio given in F table for its corresponding df_1 and df_2 at .05 level of .01 level, it is said to be significant. In such condition, we have to proceed further to test the separate differences among the two means, by applying 't' test of significance. This further procedure of testing significant difference among the two means is known as post-hoc test or post ANOVA test of difference.

To have clear understanding, go through the following working examples very carefully.

Example 1

In a study of intelligence, a group of 5 students of class IX studying each in Arts, Commerce and Science stream were selected by using random method of sample selection. An intelligence test was administered to them and the scores obtained are as under. Determine, whether the three groups differ in their level of intelligence.

Table 3.2.3

S.No.	Arts Group Intelligence scores	Comm. Group Intelligence scores	Science Group Intelligence scores
1	15	12	12
2	14	14	15
3	11	10	14
4	12	13	10
5	10	11	10

Solution: In the example k = 3 (i.e. 3 groups), n =5 (i.e. each group having 5 cases), n = 15 (i.e. the total number of units in the group)

Null hypothesis $H_0 = \mu_1 = \mu_2 = \mu_3$

i.e. the students of IX class studying in Arts, Commerce and Science stream do not differ in their level of intelligence.

Thus

Table 3.2.4

Arts Group		Commerce Group		Science Group	
X_1	X_1^2	X_2	X_2^2	X_3	X_3^2
15	225	12	144	12	144
14	196	14	196	15	225
11	121	10	100	14	196
12	144	13	169	10	100
10	100	11	121	10	100
$\sum X_1 = 62$	$\sum X_1^2 = 786$	$\sum X_2 = 60$	$\sum X_2^2 = 730$	$\sum X_2 = 61$	$\sum X_3^2 = 765$
5		5		5	
12.40		12.00		12.20	

Step 1 : Correction term

$$Cx = \frac{\sum(x)^2}{N} = \frac{\left(\sum x_1 + \sum x_2 + \sum x_3 \ldots\ldots\ldots \sum x_k\right)^2}{n_1 + n_2 + n_3 \ldots\ldots n_k} = \frac{(62+62+61)^2}{5+5+5} = \frac{(183)^2}{15}$$

Or $\quad Cx = 2232.60$

Step 2 : SS_T (Sum of squares of total) = $\sum x^2 - Cx$

Or $$= \left(\sum x_1^2 + \sum x_2^2 + \sum x_3^2 \ldots\ldots\ldots \sum x_k^2\right) - \frac{\left(\sum x\right)^2}{N}$$

$= (786+730+765) - 2232.60$

$= 2281.00 - 2232.60$

$SS_T = 48.40$

Step 3 : SS_A (Sum of squares among the groups) = $\sum \frac{\left(\sum x\right)^2}{N} - Cx$

Or $$= \frac{\left(\sum x_1\right)^2}{n_1} + \frac{\left(\sum x_2\right)^2}{n_2} + \frac{\left(\sum x_3\right)^2}{n_3} + \ldots\ldots\ldots\ldots + \frac{\left(\sum x_k\right)^2}{n_k} - Cx$$

$$= \frac{(62)^2}{5} + \frac{(60)^2}{5} + \frac{(61)^2}{5} - 2232.60$$

$= 2233.00 - 2232.60$

Or $\quad SS_A = 0.40$

Step 4 : SS_W (Sum of squares within the groups) = $SS_T - SS_A$

Or $\quad = 48.40 - 0.40$

$SS_W = 48.00$

Step 5 : MSS_A (Mean sum of squares among the groups)

$$MSS_A = \frac{SSA}{k-1} = \frac{0.40}{3-1} = \frac{0.40}{2}$$

Or $\quad MSS_A = 0.20$

Step 6 : MSS_W (Mean sum of squares within the groups)

$$= \frac{SS_W}{N-K} = \frac{48}{15-3} = \frac{48}{12}$$

$MSS_W = 4.00$

Step 7 : F Ratio = $\frac{MSS_A}{MSS_W} = \frac{0.20}{4.00} = 0.05$

Step 8 : Summary of ANOVA

Table 3.2.5 : Summary of ANOVA

Source of variance	df	SS	MSS	F Ratio
Among the Groups	(k-1) 3-1 = 2	0.40	0.20	0.05
Within the Groups	(N-k) 15-3 = 12	48.00	4.00	
Total	14			

From F table (refer to statistics book) for 2 and 12 df at .05 level, the F value is 3.59. Our calculated F value is .05, which is very low than the F value given in the table. Therefore the obtained F ratio is not significant at .05 level of significance for 2 and 12 df. Thus the null hypothesis (H_0) is accepted.

Interpretation of Results

Because null hypothesis is rejected at .05 and .01 level of significance therefore with 99% confidence it can be said that the students studying in Arts, Commerce and Science stream do not differ significantly in their level of intelligence.

Example 2

An experimenter wanted to study the relative effects of four drugs on the physical growth of rats. The experimenter took a group of 20 rats of same age group, from same species and randomly divided them into four groups, having five rats in each group. The experimenter then gave 4 drops of corresponding drug as a one doze to each rat of the concerned group. The physical growth was measured in terms of weight. After one month treatment, the gain in weight is given below. Determine if the drugs are effective for physical growth? Find out if the drugs are equally effective and determine, which drug is more effective in comparison to other one.

Table 3.2.6 : Observations (Gain in weight in ounce)

Group A (Drug P)	Group B (Drug Q)	Group C (Drug R)	Group D (Drug S)
4	9	2	7
5	10	6	7
1	9	6	4
0	6	5	2
2	6	2	7

Solution: Given k = 4, n = 5, N = 20 and Scores of 20 rats in terms of weight

Null hypothesis $H_0 = \mu_1 = \mu_2 = \mu_3$

i.e. All the four drugs are equally effective for the physical growth of the rats.

Therefore:

Table 3.2.7

	Group A		Group B		Group C		Croup D	
	X_1	X_1^2	X_2	X_2^2	X_3	X_3^2	X_4	X_4^2
	4	16	9	81	2	4	7	49
	5	25	10	100	6	36	7	49
	1	1	9	81	6	36	4	36
	0	0	6	36	5	25	2	4
	2	4	6	36	2	4	7	49
Sum	12	46	40	334	21	105	27	167
n	5		5		5		5	
Mean	2.40		8.0		4.20		5.40	

Step 1 : Correction Term Cx = $\frac{(\sum x)^2}{N} = \frac{(12+40+21+27)^2}{20} = \frac{(100)^2}{20}$

= 500.00

Step 2 : Sum of Squares of total $SS_T = \sum x^2 - Cx$

= (46+334+105+167) – 500.00

= 152

Step 3 : Sum of Squares Among groups $SS_A = \sum \frac{(\sum x)^2}{n} - Cx$

$$= \left(\frac{(12)^2}{5} + \frac{(40)^2}{5} + \frac{(21)^2}{5} + \frac{(27)^2}{5}\right) - 500.00$$

= 82.80

Step 4 : Sum of Squares Within groups $SS_W = SS_T - SS_A$

= 152 – 82.80

= 69.20

Step 5 : Summary of ANOVA

Table 3.2.8: Summary of ANOVA

Source of variance	df	SS	MSS	F Ratio
Among Groups	4-1 = 3	82.80	$\frac{82.80}{3} = 27.60$	$\frac{27.60}{4.32} = 6.39$
Within Groups (Error variance)	40-4 = 16	69.20	$\frac{69.20}{16} = 4.32$	
Total	19			

In F table $F_{.05}$ for 3 and 16 df = 3.24

and

$F_{.01}$ for 3 and 16 df = 5.24

Our obtained F ratio (6.39) is greater than the F value at .01 level of significance for 3 and 16 df. Thus the obtained F ratio is significant at .01 level of confidence. Therefore the null hypothesis is rejected at .01 level of confidence. i.e. the drugs P, Q, R, S are not equally effective for physical growth.

In the given problem it is also to be determined which drug is comparatively more effective. Thus we have to make post-hoc comparisons.

For post-hoc comparisons, we apply 't' test of significance. The common formula of 't' test is –

$$t = \frac{M_1 \sim M_2}{S.E_{DM}}$$

Where :

M_1 = Mean of first group

M_2 = Mean of second group, and

SE_{DM} = Standard Error of Difference of Means.

Here $SE_{DM} = SDW\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

Where SD_W or $\sigma_w = \sqrt{MSS_W}$

i.e. $S.D_W$ is the within groups S.D. and n_1 and n_2 are the size of the samples or groups being compared.

In the given example the means of four groups A, B, C and D are ranging from 2.40 ounce to 8.00 ounce, and the mean difference from 5.60 to 1.20. To determine the significance of the difference between any two selected means we must compute 't' ratio by dividing the given mean difference by its $S.E._{DM}$. The resulting t is then compared with the 't' value given in 't' table (Table no 2.5.1 of Unit 2) keeping in view the df of within the groups i.e. df_W. Thus in this way for four groups we have to calculate 6, 't' values as given below:

Step 6 : Standard deviation of within the groups

$$SD_W = \sqrt{MSS_W} = \sqrt{4.32}$$

= 2.08

Step 7 : Standard Error of Difference of Mean ($S.E_{DM}$)

$$S.E._{DM} = SDW\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

= 1.31

(All the groups have same size therefore the value of SE_{DM} for the two groups will remain same)

Step 8 : Comparison of the means of the various pairs of groups.

Group A vs B

$$t = \frac{M_A - M_B}{S.E_{DM}} = \frac{8.0 - 2.40}{1.31} = \frac{5.60}{1.31} = 4.28 \text{ (Significant at .01 level for 16 df).}$$

Group A vs C

$$t = \frac{4.20 - 2.40}{1.31} = \frac{1.80}{1.31} = 1.37 \text{ (Insignificant at .05 level for 16 df).}$$

Group A vs D

$$t = \frac{5.40 - 2.40}{1.31} = \frac{3.0}{1.31} \quad t = 2.29 \text{ (Significant at .05 level for 16 df).}$$

Group B vs C

$$t = \frac{8.0 - 4.90}{1.31} = \frac{3.80}{1.31} \quad t = 2.90 \text{ (Significant at .05 level for 16 df).}$$

Group B vs D

$$t = \frac{8.0 - 5.40}{1.31} = \frac{2.60}{1.31} = 1.98 \text{ (Insignificant at .05 level for 16 df).}$$

Group C vs D

$$t = \frac{5.40 - 4.20}{1.31} = \frac{1.20}{1.31} = 0.92 \text{ (Insignificant at .05 level for 16 df).}$$

Results :

Out of 6 ‘t’ values, only 3 t values are found statistically significant. Among these three, one value is found significant at .01 level, while the two values are found significant at .05 level of significance. From these ‘t’ values, it is quite clear that the group B is better in physical growth in comparison to the group A and C, similarly group D is found better in comparison to Group A. The group B & D and groups C & D are found almost equal in their physical growth.

Interpretation of the Results

Since the group B is found better in physical growth of the rats in comparison to group A at 99% confidence level and at 95% confidence level it is found better in case of group C. But the group B and D are found approximately equally good in physical growth. Therefore the drug Q and S are effective for physical growth in comparison to the drugs P and R. Further the drug Q is comparatively more effective than the other drugs P, R and S respectively.

From the forgoing illustrations you have noted that if obtained F value is not significant, it means, there is no difference either of the pairs of groups. There is no need to follow ‘t’ test. If F is found significant, then the complete procedure of analysis of variance is to specify the findings by using ‘t’ test. Therefore you have noticed that only the F value is not sufficient when it is found significant. It is to be completed when supplemented by using the ‘t’ test.

3.2.4 Steps of One Way Analysis of Variance

From the foregoing two illustration, it is clear that following steps are to be followed when we use analysis of variance.

Step 1 : Set up null hypothesis.

Step 2 : Set the raw scores in table form as shown in the two illustrations.

Step 3 : Square the individual scores of all the sets and write the same in front of the corresponding raw score.

Step 4 : Obtain all the sum of raw scores and the squares of raw scores. Write them at the end of each column.

Step 5 : Obtain grand sums of raw scores as and square of raw square

as $\sum x^2$

Step 6 : Calculate correction term by using the formula

$$\text{Cx} = \frac{\sum x^2}{N} \quad \text{Or} \quad \text{Cx} = \frac{\left(\sum x_1 + \sum x_2 + \sum x_3 \ldots\ldots\ldots\ldots + \sum x_k\right)^2}{n_1 + n_2 + n_3 + \ldots\ldots\ldots n_k}$$

Step 7 : Calculate sum of squares i.e. SS_T by using the formula-

$$SS_T = \sum x^2 - Cx$$

Step 8 : Calculate sum of squares among the groups i.e. SS_A by using the formula-

$$SS_A = \frac{\sum x^2}{n} - Cx$$

$$\text{Or } SS_A = \frac{\left(\sum x_1^2\right)^2}{n_1} + \frac{\left(\sum x_2^2\right)^2}{n_2} + \frac{\left(\sum x_3^2\right)^2}{n_3} + \ldots\ldots\ldots\ldots + \frac{\left(\sum x_k^2\right)^2}{n_k} - Cx$$

Step 9 : Calculate sum of squares within the groups i.e. SSw by using the formula

$$SS_W = SS_T - SS_A$$

Step 10 : Calculate the degrees of freedom as

greater degree of freedom $df_1 = k - 1$ (where k is number of groups)

Smaller degree of freedom $df_2 = N\text{-}k$ (where N is the total number in the group)

Step 11 : Find the value of Mean sum of squares of two variances as-

Mean sum of squares between the group $MSS_A = \frac{SS_A}{k-1}$

Mean sum of squares within the groups $MSS_W = \frac{SS_W}{N-K}$

Step 12 : Prepare summary table of analysis of variance as shown in 3.2.5 or 3.2.8.

Step 13 : Evaluate obtained F Ratio with the F ratio value given in F table (Table no. 3.3.1) keeping in mind df_1 and df_2.

Step 14 : Retain or Reject the Null Hypothesis framed as in step no-I.

Step 15 : If F ratio is found insignificant and null hypothesis is retained, stop further calculation, and interpret the results accordingly. If F ratio is found significant and null hypothesis is rejected, go for further calculations and use post-hoc comparison, find the t values and interpret the results accordingly.

3.2.5 Assumptions Underlying the Analysis of Variance

The method of analysis of variance has a number of assumption. The failure of the observations or data to satisfy these assumptions, leads to the invalid inferences. The following are the main assumptions of analysis of variance.

The distribution of the dependent variable in the population under study is normal.

There exists homogeneity of variance i.e. the variance in the different sets of scores do not differ beyond chance, in other words $\sigma_1 = \sigma_2 = \sigma_3 = ... = \sigma_k$.

The samples of different groups are selected from the population by using random method of sample selection.

There is no significant difference in the means of various samples or groups taken from a population.

3.2.6 Relationship between 'F' test and 't' test

The F test and t test are complementary to each other, because-

't' is followed when 'F' value is significant for the specification of inferences.

'F' test is followed, when 't' value is not found significant. Because within groups variance is not evaluated by 't' test. It evaluate only the difference between variance.

There is a fixed relation between 't' and 'F'. the F is the square of 't', while 't' is a square root of F.

F = t^2 $or\ t = \sqrt{F}$

3.2.7 Merits or Advantages of Analysis of Variance

The analysis of variance technique has the following advantages:

- It is the improved technique over the 't' test or 'z' test, it evaluates both types of variance 'between' and 'within'.
- This technique is used for ascertaining the difference among several groups or treatments at a time. It is an economical device.
- It can involve more than one variable in studying their main effects and interaction effects.
- In some of the experimental design e.g. simple random design and levels X treatment designs are based on one way analysis of variance.
- If 't' is not significant, F test must be followed, to analyse the difference between two means.

3.2.8 Demerits or Limitations of Analysis of Variance

The analysis of variance technique has following limitations also:

- We have seen that analysis of variance techniques is based on certain assumptions e.g. normality and homogeneity of the variances among the groups. The departure of the data from these assumptions may effect adversely on the inferences.
- The F value provides global findings of difference among groups, but it can not specify the inference. Therefore, for complete analysis of variance, the 't' test is followed for specifying the statistical inference.

- It is time consuming process and requires the knowledge and skills of arithmetical operations as well the high vision for interpretations of the results.
- For the use of 'F' test, the statistical table of 'F' value is essential without it results can not we interpreted.

3.3 F RATIO TABLE AND THE PROCEDURE TO USE

The significance of difference between two means is analysed by using 't' test or 'z' test as it has been discussed in earlier unit 2. The calculated or obtained t value is evaluated with the help of 't' distribution table. With df at .05 and .01 levels of significance. The df is the main base to locate the 't' values at different levels of significance given in the table.

In a similar way the calculated F ratio value is evaluated by using F table (refer to statistics book) by considering the degree of freedom between the groups and within the groups.

You observe the F table, carefully, you will find there are rows and columns. In the first row there are the degrees of freedom for larges variance i.e. for greater mean squares or between the variance. The first column of the table has also degree of freedom of smaller mean squares or the variance within the groups. Along with these two degree of freedom the F ratio values are given at .05 and .01 level of significance. The normal print values is the values at .05 level and the bold or dark print values are at .01 level of significance.

In the first illustrated example in the summary table the F ratio value is 0.05, df_1, is 2 and df_2 is 12. For evaluating the obtained F value with F value given in table df_1 = 2 is for row and df2 is for column. In the column you will find 12, proceed horizontally or row wise and stop in column 2, you will find F values 3.88, which is for .05 level of confidence and 6.93 (in dark bold print) which is meant for .01 level of confidence. Our calculated F value .05 is much less than these two values. Hence the F ratio is not significant also at .05 level. Thus the null hypothesis is retained.

Self Assessment Questions

1) State the assumptions of ANOVA.

....

....

....

....

2) What happens when these assumptions are violated?

....

....

....

....

3) Compare the 'F Ratio' test and 't Ratio' test in terms of their relative merits and demerits.

....

....

....

....

4) What is the mathematical relationship between F and t.

....

....

....

....

5) What is the relationship between S.D. and Mean sum of squares within the groups?

....

....

....

....

6) When the post ANOVA test of difference is applied?

....

....

....

....

7) How many degree of freedom are associated with the variation in the data for- A comparison of four means for independent samples each containing 10 cases?

....

....

....

....

8) A comparison of three groups selected independently each containing 15 units.

....

....

....

....

3.4 LET US SUM UP

Analysis of variance is used to test the significance of the difference between the means of a number of different populations say two or more than two.

Analysis of variance deals with variance rather to deal with means and their standard error of the difference exist between the means.

The variance is the most important measure of variability of a group. It is simply the square of S.D. of the group i.e. v = σ^2

The problem of testing the significance of the differences between the number of means results from experiments designed to study the variation in a dependent variable with variation in independent variable.

Analysis of variance is used when difference in the means of two or more groups is found insignificant.

There is a fixed relationship between 't' ratio and 'F' ratio. The relationship can be expressed as $F = t^2$ or $t = \sqrt{F}$.

While determining the significance of calculated or obtained ratio, we consider two types of degrees of freedom. One greater i.e. degree of freedom between the groups and second smaller i.e. degree of freedom within the groups.

3.5 UNIT END QUESTIONS

1) The four groups are given four treatments, each group consists of 5 subjects. At the end of treatment a test is administered, the obtained scores are given in the following table. Test significance of difference among four treatments.

Scores of the treatment

Group – A	Group – B	Group – C	Group – D
X1	X2	X3	X4
14	19	12	17
15	20	16	17
11	19	16	14
10	16	15	12
12	16	12	17

2) A Test Anxiety test was given to three groups of students of X class, classified as high achievers, average achievers and low achievers. The scores obtained on the test are shown below. Are the three groups differ in their test anxiety.

High-achievers	Average achievers	Low achievers
15	19	12
14	20	14
11	16	12
12	19	12

3) Apply ANOVA on the following sets of scores. Interpret your results.

Set-I	Set-II	Set-III
10	3	10
7	3	11
6	3	10
10	3	5
4	3	6
3	3	8
2	3	9
1	3	12
8	3	9
9	3	10

4) Summary of analysis of variance is given below:

Source of variance	Df	SS	MSS	F
Between sets	2	180	90.00	17.11
Within sets	27	142	5.26	
Total	29			

Interpret the result obtained.

Note Table F values are

$F_{.05}$ for 2 and 27 df = 3.35

$F_{.01}$ for 2 and 27 df = 5.49

5) Given the following statistics of the two groups obtained on a verbal reasoning test:

Group	N	M	σ
Boys	95	29.20	11.60
Girls	83	30.90	7.80

Calculate:

't' ratio for the two groups.

'F' ratio for the two groups.

What should be the degree of freedom for 't' ratio.

What should be the degrees of freedom for 'F' ratio.

Interpret the results obtained on 't' ratio and 'F' ratio.

6) Why it is necessary to fulfill the assumptions of 'F' test, before to apply analysis of variance.

7) Why the ‘F’ ratio test and ‘t’ ratio tests are complementary to each other.

8) What should be the various problems of psychology and education. Where the ANOVA can be used successfully.

3.6 SUGGESTED READINGS

Aggarwal, Y.P. (1990). *Statistical Methods – Concept, Applications, and Computation*. New Delhi : Sterling Publishers Pvt. Ltd.

Ferguson, G.A. (1974). *Statistical Analysis in Psychology and Education.* New York : McGraw Hill Book Co.

Garret, H.E. & Woodwarth, R.S. (1969). *Statistics in Psychology and Education.* Bombay : Vakils, Feffer & Simons Pvt. Ltd.

Guilford, J.P. & Benjamin, F. (1973). *Fundamental Statistics in Psychology and Education.* New York : McGraw Hill Book Co.

Srivastava, A.B.L. & Sharma, K.K. (1974). *Elementary Statistics in Psychology and Education.* New Delhi : Sterling Publishers Pvt. Ltd.

Walker, H.M. & Lev. J. (1965). *Statistical Inference.* Calcutta : Oxford & I.B.H. Publishing Co.

UNIT 4 TWO WAY ANALYSIS OF VARIANCE

Structure

4.0 INTRODUCTION

In the preceding unit 3 we have learned about the one way analysis of variance. In this technique, the effect of one independent or one type of treatment was studied on single dependent variable, by taking number of groups from a population or from different population heaving different characteristics. Generally, in one way analysis of variance simple random design is used.

Now, suppose we want to study the effect of two independent variables on a single dependent variable. Further suppose our aim is to study the independent effects of the independent variables as well as their combined or joint effect on the dependent variable. For example a medicine company has developed two types of drugs to get relief from smoking habit. The company wants to know:

1) The independent effect of drug A on smoking behaviour,

2) The independent effect of drug B on smoking behaviour, and

3) The joint or interactional effect of drug A and B i.e. A x B on the smoking behaviour.

Take another example, in a field experiment, a psychologist wants to study effect of type of families on the cognitive development of the children of the age group 3 to 5+ years of age in relation to their sex.

In this field experiment there are two independent variables viz. Type of Family and gender of the Children. The dependent variable is Cognitive Development.

Further the type of family variable has two levels i.e. joint families and nuclear families.

Similarly the gender variable has also two levels viz. boys and girls.

The experimenter wants to study the independent effects of type of family (Joint vs Nuclear) gender (Boys vs Girls) and the interactional effect i.e. joint effect of type of family and gender on the dependent variable viz. Cognitive Development.

Such type of studies related to field experiments or real experiments are known as factorial design of 2 × 2 which indicates there are two independent variables each having two levels.

Like wise there are several situations in which the effect of two or more than two independent variable is studied on a single dependent variable.

In such experimental studies, the one way analysis of variance is not applicable. We have to use two, three or four way of analysis of variance, which depends upon the number of independent variables and their number of levels.

4.1 OBJECTIVES

After completing this unit, you will be able to:

- Define two way analysis of variance;
- Use analysis of variance vertically or column wise and horizontally or row wise;
- Explain the independent effects of two or more than two variables having each two or more than two levels;
- Explain the term interaction effect;
- Analyse the interaction effect of two or more than two variables;
- Differentiate between one way analysis of variance and two way analysis of variance;
- Analyse problems related to field experiments and true experiments where factorial designs are used;
- Explain the interactional effect of two variables on dependent variables; and
- Explain variables graphically.

4.2 TWO WAY ANALYSIS OF VARIANCE

In two way analysis of variance, usually the two independent variables are taken simultaneously. It has two main effects and one interactional or joint effect on dependent variable. In such condition we have to use analysis of variance in two way i.e. vertically as well as horizontally or we have to use ANOVA, column and row wise. To give an example , suppose you are interested to study the intelligence i.e. I.Q. level of boys and girls studying in VIII class in relation to their level of socio economic status (S.E.S.). in such condition you have the following 3 x 2 design.(Refer to table 4.2.1)

Table 4.2.1: SES, Intelligence and Gender factors

Groups	Levels of S.E.S.			
	High	Average	Low	Total
Boys	M_{HB}	M_{AB}	M_{LB}	M_B
Girls	M_{HG}	M_{AG}	M_{LG}	M_G
Total	M_H	M_A	M_L	M

In the table above,

M : Mean of intelligence scores.

M_{HB}, M_{AB}, & M_{LB} : Mean of intelligence scores of boys belonging to different levels of S.E.S. i.e. High, Average & Low respectively.

M_{HG}, M_{AG}, & M_{LG} : Mean of intelligence scores of girls belonging to different levels of S.E.S. respectively.

M_H, M_A, & M_L : Mean of the intelligence scores of students belonging to different levels of S.E.S. respectively.

M_B, M_G : Mean of the intelligence scores of boys and girls respectively.

From the above 3 x 2 contingency table, it is clear, first you have to study the significant difference in the means column wise or vertically, i.e. to compare the intelligence level of the students belonging to different categories of socio-economic status (High, Average and Low).

Second you have to study the significant difference in the means row wise or horizontally, i.e. to compose the intelligence level of the boys and girls.

Then you have to study the interactional or joint effect of sex and socio-economic status on intelligence level i.e. we have to compare the significant difference in the cell means of columns and rows.

Obviously, you have more than two groups, and to study the independent as well as interaction effect of the two variables viz. socio-economic status and sex on dependent variable viz. intelligence in terms of I.Q., you have to use two way analysis of variance i.e. to apply analysis of variance column and row wise.

Therefore, in two way analysis of variance technique, the following type of effects are to be tested:

- Significance of the effect of A variable on D.V.
- Significance of the effect of B variable on A.V.
- Significance of the interaction effect of A x B variables on D.V.

In two way analysis of variance, the format of summary table after applying the analysis of variance is as under-

Table 4.2.2: Summary of two way ANOVA

Source of variance	df	SS	MSS	F Ratio
Among the groups				
Between the group A	$k_a - 1$	SS_A	MSS_A	$F_1 = \frac{MSS_A}{MSS_W}$
Between the Group B	$k_b - 1$	SS_B	MSS_B	$F_2 = \frac{MSS_B}{MSS_W}$
Interrelation A x B	$(k_a - 1)(k_b - 1)$	SS_{AxB}	MSS_{AxB}	$F_3 = \frac{MSS_{A\times B}}{MSS_W}$
Within the Groups (Error variance)	$N - k_a - k_b$			
Total	N – 1			

For interpretation of the obtained F ratios, we have to evaluate each F ratio value with the F ratio given in F table (refer to statistics book) keeping in view the corresponding grater and smaller df and the level of confidence. There may be two possibilities.

All the obtained F ratios may be found insignificant even at .05 level. This shows that there is no independent (i.e. individual) as well as interaction (i.e. joint) effect of the two independent variables on dependent variable. Hence null hypothesis will retain. There is no need to do further calculations.

All the three obtained F ratio's may be found significant either at .05 level of significance or at .01 level of significance. This shows that there is a significant independent (i.e. individual) as well as interactional (i.e. joint) effect of the independent variables on the dependent variable. Therefore the null hypothesis is rejected. In such condition if the two independent variables have more than two levels i.e. three or four, we have to go for further calculations and use post-hoc comparisons by finding out various 't' values by pairing the groups.

Similarly the significant interactional effect will also be studied further by applying 't' test of significance or by applying graphical method.

At least one or two obtained F ratio will be found significant either at .05 level of significance or at .01 level of significance. Thus the null hypothesis may partially be retained. In such condition too we have to do further calculations, by making post-hoc comparisons and use 't' test of significance, if the independent variables have more than two levels.

For more clarification, go through the following illustration carefully.

Example 1

A researcher wanted to study the effect of anxiety and types of personality (Extroverts and Introverts) on the academic achievement of the undergraduate students. For the purpose, he has taken a sample of 20 undergraduates by using random method of sample selection. He administered related test and found following observations in relation to the academic achievement of the students.

Level of Anxiety

<table>
<tr><td></td><td>Groups</td><td>High anxiety</td><td>Low anxiety</td></tr>
<tr><td rowspan="2">Type of Personality</td><td>Extroverts</td><td>12
13
14
15
14</td><td>14
14
13
15
15</td></tr>
<tr><td>Introverts</td><td>14
16
16
16
15</td><td>11
10
12
12
16</td></tr>
</table>

Determine the independent as well as interactional effect of anxiety and types of personality on the academic achievement of the undergraduates.

Solutions:

Given

Two independent variables

a) type of personality having two levels viz. extroverts and introverts

b) Anxiety it has also two level viz. high anxiety and low anxiety.

Dependent variable scores

Academic achievement scores.

Number of groups i.e. k = 4.

Number of units in each cell i.e. n = 5.

Total no. of units i.e. N = 20.

To find out :

Independent effect of type of personality and anxiety on the academic achievement of the students.

Interactional i.e. joint effect of anxiety and type of personality on academic achievement of the students.

Therefore

H_0 : "There is no significant effect of types of personality and level of anxiety on academic achievement."

For convenience, the given 2 x 2 table is rearranged as under:

Table 4.2.3

S.N.	Extroverts				Introverts			
	High Anxiety		Low Anxiety		High Anxiety		Low Anxiety	
	X_1	X_1^2	X_2	X_2^2	X_3	X_3^2	X_4	X_4^2
1	12	144	14	196	14	196	11	121
2	13	169	14	196	16	256	10	100
3	14	196	13	169	16	256	12	144
4	15	225	15	225	16	256	12	144
5	14	196	15	225	15	225	16	256
Sums	68	930	71	1011	77	1189	61	765
N	5		5		5		5	
M	13.60		14.20		15.40		12.20	

Step 1 : Correction Term Cx = $\frac{\left(\sum x^2\right)}{N}$

$= \frac{(68+71+77+61)^2}{20} = \frac{(277)^2}{20}$

= 3836.45

Step 2 : Sum of Squares of Total $SS_T = \sum x^2 - Cx$

$= 930+1011+1189+765 - 3836.45$

$= 58.55$

Step 3 : Sum of Squares Among the Groups

$$SS_A = \sum \frac{(\sum x)^2}{n} - Cx$$

$= \frac{(68)^2}{5} + \frac{(71)^2}{5} + \frac{(77)^2}{5} + \frac{(61)^2}{5} = 3836.45$

= 26.55

Step 4 : Sum of squares Between the A Groups (i.e. between types of personality)

$$SS_{BTP} = \frac{\left(\sum x_1 + \sum x_2\right)^2}{n_1 + n_2} + \frac{\left(\sum x_3 + \sum x_4\right)^2}{n_3 + n_4} - Cx$$

$= 3836.50 - 3836.45$

$= .05$

Step 5 : Sum of squares Between the B Groups (i.e. Between level of Anxiety)

$$SS_{Anx} = \frac{\left(\sum x_1 + \sum x_2\right)^2}{n_1 + n_2} + \frac{\left(\sum x_3 + \sum x_4\right)^2}{n_3 + n_4} - Cx$$

$$= \frac{(68+77)^2}{5+5} + \frac{(71+61)^2}{5+5} - 3836.45$$

$= 8.45$

Step 6 : Sum of squares of Interaction

$$SS_{AxB} = SS_A - SS_{BTP} - SS_{BAnx}$$

i.e. SS_{AxB} = Sum of squares Among the Groups – Sum of Squares Between Type of Personality – Sum of Squares Between Anxiety Levels.

$$SS_{AxB} = 26.55 - 0.05 - 8.45$$

$$= 18.05$$

Step 7 : Sum of Squares Within the Groups

$$SS_W = SS_T - SS_A - SS_B$$

$$= 58.55 - 26.55$$

$$= 32.00$$

Step 8 : Preparation of Result of Summary Table

Table 4.2.4 : Summary of Two-way ANOVA

Source of variance	df	Sum of Squares (SS)	Mean SS (MSS)	F Ratio
Among the Groups	(k-1) (4-1=3)	(26.55)	$\frac{SS_A}{df} = \frac{26.55}{3} = 8.85$	$\frac{8.85}{2} = 4.425$
Between the Groups- SS_{B_1} (Types of personality)	(k_1-1) 2 – 1 = 1	0.05	$\frac{SS_{B_1}}{df} = \frac{.05}{1} = .05$	$\frac{.05}{2} = .025$
SS_{B_2} (Anxiety levels)	(k_2-1) 2 – 1 = 1	8.45	$\frac{SS_{B_2}}{df} = \frac{8.45}{1} = 8.45$	$\frac{8.45}{2} = 4.225$
SS_{AxB}	$(k_1-1)(k_2-1)$ 1 x 1 = 1	18.05	$\frac{SS_{B_1} \times SS_{B_2}}{df} = \frac{18.05}{1} = 18.05$	$\frac{18.05}{2} = 9.025$
Within the Groups	(N-k) 20 – 4 = 16	32.00	$\frac{SS_W}{df} = \frac{32}{16} = 2$	
Total	19			

In the F table (refer to statistics book) for 1 and 6 df, the F value at .01 and .05 level are 8.86 and 4.60 respectively.

Our calculated F values for type of personality and anxiety are smaller than the table F value 4.60.

Therefore the obtained F ratio values are not significant even at .05 level of significance. Hence the null hypotheses is in relation to Type of Personality and Anxiety are retained.

In case of interaction effect the obtained F ratio value 9.025 is found higher than the F value given in table at .01 level of significance. Thus the F for interaction effect is significant at .01 level. Hence, null hypothesis for interaction effect is rejected.

Interpretation of the Results

Since our null hypotheses are accepted at .05 and .01 level of significance, for type of personality, therefore it can be said that there is no independent as well as interactional effect of Types of Personality and levels of Anxiety on the academic achievement of the students. In other words it can be said that the students who are either Extroverts or Introverts are equally good in their academic performance.

Similarly, the anxiety level of the students do not cause any significant variation in the academic achievement of the students.

But the students having different type of personality and have different level of anxiety, their academic achievement varies in 99% cases. From the mean values in the table 4.2.3 it is evident that the students who are Extroverts and have low level of anxiety are comparatively good in their academic achievement (M = 14.20).

In the case of Introverts those who have high level of anxiety are better in their academic achievement (M = 15.40) in comparison to others.

Example 2

In a study, effect of intelligence and sex on the mathematical creativity a group of 40 students (20 boys and 20 girls) was selected from a population of high school going students by using random method of sample selection. A test of intelligence and mathematics creativity was administrated to them. The observations obtained are given below. Determine the independent as well as interactional effect of sex and Intelligence on the mathematical creativity of the high school going students.

Table 4.2.5: Observations obtained on the mathematical creativity test

Groups	Boys	Girls
High Intelligent	15 15 15 12 13 15 16 16 16 20	14 13 13 15 15 13 13 14 15 14
Low Intelligent	15 14 12 13 15 14 15 14 13 12	10 12 10 13 13 10 11 12 10 10
Total units	20	20

Solution:

Given :

Two independent variables A- Sex, B- Intelligence. Each having 2 levels.

Dependent variable : Mathematical Creativity

Number of Groups k = 4

Number of cases in each group n = 10

Total number of units in the group N = 40

To find out : i) Independent effect of intelligence and sex on mathematical creativity.

ii) Interactional effect of intelligence and sex on mathematical creativity.

H_0 : There is no significant independent as well as interactional effect of Intelligence and Sex on the mathematical creativity of the students.

Therefore.

Table 4.2.6

Groups	Boys (A1)				Girls (A2)			
S.No.	High Intelligence (B1)		Low Intelligence (B2)		High Intelligence (B1)		Low Intelligence (B2)	
	X1	X2	X1	X2	X1	X2	X1	X2
1	15	225	15	225	14	196	10	100
2	15	225	14	196	13	169	12	144
3	15	225	12	144	13	169	10	100
4	12	144	13	169	15	225	13	169
5	13	169	15	225	15	225	13	169
6	15	225	14	196	13	169	10	100
7	16	256	15	225	13	169	11	121
8	16	256	14	196	14	196	12	144
9	16	256	13	169	15	225	10	100
10	20	400	12	144	14	196	10	100
Sum	153	2381	137	1889	139	1939	111	1247
n	10		10		10		10	
Mean	15.30		13.70		13.90		11.10	

Step 1 : Correction Term = Cx = $\frac{\sum(x)^2}{N} = \frac{(153+137+139+111)^2}{40}$

= 7290.00

Step 2 : Sum of Squares of total SS_T = $\sum x^2 - Cx$

= (2381+1889+1939+1247) – 7290

= 166.00

Step 3 : Sum of Squares Among groups SS_A = $\sum\frac{\sum(x)^2}{N} - Cx$

$$= \left(\frac{(153)^2}{10} + \frac{(137)^2}{10} + \frac{(139)^2}{10} + \frac{(111)^2}{10}\right) - 7290.00$$

= 92

Step 4 : Sum of squares Between the Groups (Sex)

$$SS_{BSex} = \frac{(\sum x_1 + \sum x_2)^2}{n_1 + n_2} + \frac{(\sum x_3 + \sum x_4)^2}{n_3 + n_4} - Cx$$

$$= \frac{(153+137)^2}{20} + \frac{(139+111)^2}{20} - 7290$$

= 40.00

Step 5 : Sum of squares Between the Groups (Intelligence)

$$SS_{BInt} = \frac{(\sum x_1 + \sum x_2)^2}{n_1 + n_2} + \frac{(\sum x_3 + \sum x_4)^2}{n_3 + n_4} - Cx$$

$$= \frac{(153+139)^2}{10+10} + \frac{(137+111)^2}{10+10} - 7290.00$$

= 48.40

Step 6 : Sum of squares Between the Interactions (Sex x Intelligence)

$$SS_{BSex \times Int} = SS_A - SS_{BSex} - SS_{BInt}$$

$$= 3.60$$

Step 7 : Sum of Squares within the Groups

$$SS_W = SS_T - SS_A$$

$$= 166 - 92$$

$$= 74.00$$

Step 8 : Preparation of Summary Table / Result table

Table 4.2.7 : Summary of Analysis of Variance

Source of variance	df	Sum of Squares SS	Mean SS MSS	F Ratio
1) Among the Groups	(k-1) 4-1=3	(92)	(30.67)	(14.88)
2) Between the Groups i) $SS_{B_{Sex}}$	(k_1-1) 2-1=1	40.00	40.00	19.42
ii) $SS_{B_{Int.}}$	(k_2-1) 2-1=1	48.40	48.40	23.49
iii) $SS_{B_{Sex \times Int.}}$	1x1=1	3.60	3.60	1.75
3.Within the Groups (Error variance)	(N-k) 40-4=36	74.00	2.06	
Total	39			

From F table, the value of $F_{.05}$ for 1 and 36 df = 4.12 and $F_{.01}$ for 1 and 36 df = 7.42

Interpretation of the Results:

Independent Effects

Sex : From the ANOVA summary table the F ratio value for Sex is found 19.42, which is high in comparison to the F value given in F table for 1 and 36 df. Therefore F ratio for Sex variable is found significant at .01 level. Hence null hypothesis is rejected. In conclusion it can be said that in 99% cases, the boys are high in mathematical creativity in comparison to the girls. There are only 1% chance that the girls are better in mathematical creativity than the boys.

Intelligence: From the ANOVA summary table the F ratio value for intelligence is found 23.49, which is also significant at .01 level for 1 and 36 df. Thus the null hypothesis is rejected at .01 level of confidence.

Therefore, in 99% cases the high intelligent high school going students are high in their mathematical creativity in comparison to the low intelligent students. Only in 1 case out of 100, the low intelligent high school going students are high in mathematical creativity.

Interactional Effect

From the ANOVA summary table it is evident that the F ratio for interactional effect is found insignificant even at .05 level of significance for 1 and 36 df. Thus the null hypothesis is accepted.

Therefore, the joint effect of sex and intelligence do not cause any significant variation in the scores of mathematical creativity. In other words both boys and girls who are high in their intelligence are equally good in their mathematical creativity.

Similarly the low intelligent boys and girls also do not differ in their mathematical creativity. In the group of boys the high intelligent and low intelligent high school going students also do not differ in their mathematical creativity. Similarly in the group of girls, the high intelligent and low intelligent girls are also do not differ significantly in their mathematical creativity. This fact is also confirmed from the following Figure A and B.

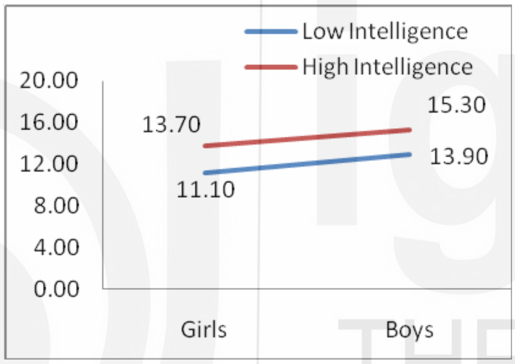

Fig. 4.2.1(A)

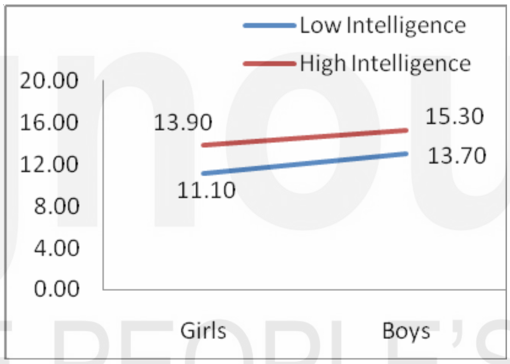

Fig. 4.2.1(B)

The two figures 4.2.1 (A) and 4.2.1 (B) both are showing two parallel lines. Which indicates that there is no interaction effect of sex and intelligence on the mathematical creativity of the high school going students.

Self Assessment Questions

1) What is the difference between one way analysis of variance and two way analysis of variance?

....

....

....

....

2) When we use two way analysis of variance?

....

....

....

....

3) In two way analysis of variance how many effects are tested.

....

....

....

4) What indicates $K_{(a)}$, $K_{(b)}$ and $K_{(c)}$

....

....

....

....

5) What is meant by df_1 and df_2?

....

....

....

....

6) In what way we decide the significance of F ratio obtained in relation to various effects?

....

....

....

....

7) What do you mean by

2 x 2 Level design

3 x 3 Level design

2 x 4 Level design

3 x 3 Level design

....

....

....

....

4.3 INTERACTIONAL EFFECT

In the foregoing discussion we have frequently used the term "interaction" or "interactional effect" therefore it is essential to clarify the same.

In the two way analysis of variance, the consideration and interpretation of the interaction of variables or factors become important. Without considering the interaction between the different variables in a study, there is no use of two way or three way analysis of variance.

The interactions may be between two or more than two independent variables and its effect is measured on the dependent variable or the criterion variable. The need to know interaction effect on criterion variable or dependent variable is to know the combined effect of two or more than two independent variables on the criterion variable.

The reason, suppose there are two independent variables A & B and each has their own significance to create high variations in the criterion or dependent variable, but their joint or combined effect may cause very high, or low or their nullified effect on the dependent variable or criterion variable.

To have more clear idea let us suppose we have two types of fertilizers e.g. Urea and Phosphate. These have their own importance in the growth of the corps independently. But when we use the two chemical fertilizers in combined way with proper ratio, it might be possible, the growth of crops may be increased tremendously. Or it might be possible, the growth of crops may go down.

In the field of psychology and education suppose a treatment or a method of teaching A has its own significance to increase the level of achievement in a school subject, similarly the treatment B or a teaching method B is also good for encouraging results in academic achievement. But when we use the two methods of teaching jointly or give two treatments in combination we may get more encouraging results in academic achievement, or we do not have any significant effect on the achievement of the learners.

In the illustration-2, presented in this unit, compare the mean values, which we have shown in each cell in table 4.2.6. for convenience we are taking the same values here and presenting in the following 2 x 2 table.

Table 4.3.1

Intelligence	Boys M_1	Girls M_2	Total Mean
High	15.30	13.90	14.60
Low	13.70	11.10	12.40
Total Mean	14.50	12.50	13.50

In the above table, if we compare the total mean of first and second column, it is quite clear that there is a difference in the mean values of boys and girls and the higher mean is in the favour of boys. This is an independent effect of sex on mathematical creativity.

Similarly if we compare the total means of two rows we find, there is a difference in the means of high intelligent and low intelligent students and higher mean is in favour of the high intelligent group. It is actually the independent effect of intelligence on mathematical creativity.

Further in the above table 4.3.1, sex effects for boys and girls are

(14.50-13.50) = 1 and (12.50 – 13.50) = -1 respectively. If we subtract the first effect 1, from all averages in the first row and add 1 to all the averages in the second row, we have the following table:

Table 4.3.2: Sex factor

Groups	Boys	Girls	Total M
High Intelligence	14.30	12.90	13.60
Low Intelligence	14.70	12.10	13.40
Total Mean	14.50	12.50	13.50

Similarly in table 4.3.1 we subtract 1 from first column and add to the second column we have the following table:

Table 4.3.3: Sex factor

Groups	Boys	Girls	Total M
High Intelligence	14.30	12.90	14.60
Low Intelligence	14.70	12.10	12.40
Total M	13.50	13.50	13.50

Table 4.3.2 and table 4.3.3 give the intersectional resultant average, which show the direction of interaction and also indicates that there is no interaction effect of the A and B independent variable on dependent variable. In such condition if we plot the graph between the two independent variables we have approximately two parallel lines, as we have seen in the graphical presentation (see fig. 4.2.1 A and 4.2.1 B) respectively.

If there is a significant interactional effect of the two or more independent variables on the dependent variables; in such condition the graphical representation of the interactional effect will show two lines which are interacting at a point. For example, in example 1 the interactional effect of type of personality and anxiety is found significant at .01 level. If we draw the graph for interaction effect of Type of Personality and Level of Anxiety by considering the mean values of academic achievement, the obtained graph will be as under. (table 4.3.4. and graphs figures 4.3.1. A and B)

Table 4.3.4: the mean values of Extroverts and Introverts having high and low level of anxiety

Groups	Extroverts M1	Introverts M2	Total Mean
High Anxiety	13.60	15.40	14.50
Low Anxiety	14.20	12.20	13.20
Total Mean	13.90	13.80	13.85

(Mean values from table 4.2.3)

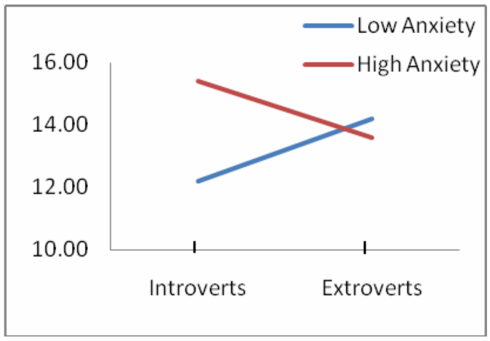

Fig. 4.3.1 (A)

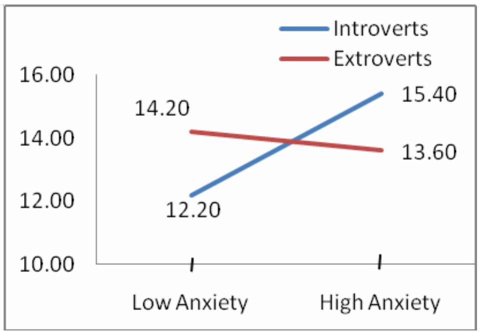

Fig. 4.3.1 (B)

4.4 MERITS AND DEMERITS OF TWO WAY ANOVA

4.4.1 Merits of Two Way Analysis of Variance

The following are the advantages of two way analysis of variance-

- This technique is used to analyse two types of effects viz. main effects and Interaction Effects.
- More than two factors effects are analysed by this technique.
- For analysing the data obtained on the basis of factorial designs, this technique is used.
- This technique is used to analyse the data for complex experimental studies.

4.4.2 Demerits or Limitations of Two Way ANOVA

The following limitations are found in this technique:

- When there are more than two classification of a factor or factors of study. F ratio value provides global picture of difference among the main treatment effects. The inference can be specified by using 't' test in case when F ratio is found significant for a treatment.
- This technique also follows the assumptions on which one way analysis of variance is based. If these assumptions are not fulfilled, the use of this technique may give us spurious results.
- This technique is difficult and time consuming.
- As the number of factors are increased in a study, the complexity of analysis in increased and interpretation of results become difficult.
- This technique requires high level arithmetical and calculative ability. Similarly it also requires high level of imaginative and logical ability to interpret the obtained results.

4.5 LET US SUM UP

The two way analysis of variance is a very important parametric technique of inferential statistics. It helps in taking concrete decisions about the effect of various treatments on criterion or dependent variable independently and jointly.

The independent effect of a variable on treatment means the direct or isolate effect of it on the dependent or criterion variable.

The interactional effect means joint effect of the two or more variables acting together on a dependent or criterion variables.

In case of insignificant interactional effect of the two or more predictors or independent variables on a criterion or dependable variable. The graphical representation will show the parallel lines, when the interactional effect is significant, its graphical representation will show the two crossed lines.

The two way analysis variance is a useful technique in experimental psychology as well as in experimental education especially in the field of teaching and learning. It is frequently used in field experiments or true experiments, when we use factorial designs specifically.

Interaction variance might be more reasonably expected in a combination of teacher and instruction method, of kind of task and method of attack by the learner, and of kind of reward when combined with a certain condition of motivation.

4.6 UNIT END QUESTIONS

1) What do you mean by Two-way Analysis of Variance.?
2) What is the difference between one way and two way ANOVA?
3) Indicate the graphical presentation of interaction effects?
4) Highlight the advantages and limitations of two way analysis of variance.
5) From the following hypothetical data, (Table below) determine-Which teaching method is effective than others. Also, Which teacher is contributing effectively in the learning outputs of the learners.
6) How far the joint effect of teaching method and the teacher is contributing in the learning performance of the students.

Teaching Methods

Teacher	A	B	C
T_1	10 7 6 10 4	3 3 3 3 3	10 11 10 5 6
T_2	3 1 8 9 2	3 3 3 3 3	8 9 12 9 10

Four groups of 8 students each having an equal number of boys and girls were selected randomly and assigned to different four conditions of an experiment. Test main effects due to conditions and sex and the interaction of the two conditions

Graphs	I	II	III	IV
Boys	7	9	12	12
	0	4	6	14
	5	5	10	9
	8	6	6	5
Girls	3	4	3	6
	3	7	7	7
	2	5	4	6
	0	2	6	5

7) In 4 × 3 factorial design 5 subjects are assigned randomly in each graph of 12 cells.

The following data obtained at the end of the experiment

Level of Intelligence	Method of Teaching			
	M1	M2	M3	M4
High (L_1)	6	8	7	9
	2	3	6	6
	4	7	9	8
	2	5	8	8
	6	2	5	9
Average (L_2)	4	6	9	7
	1	6	4	8
	5	2	8	4
	2	3	4	7
	3	6	8	4
Low (L_3)	4	3	6	6
	2	1	4	5
	1	1	3	7
	1	2	8	9
	2	3	4	8

Test the significance difference of difference of main effects and interaction effects.

4.7 SUGGESTED READINGS

Aggarwal, Y.P. (1990). *Statistical Methods-Concept, Applications, and Computation*. New Delhi : Sterling Publishers Pvt. Ltd.

Ferguson, G.A. (1974). *Statistical Analysis in Psychology and Education.* New York : McGraw Hill Book Co.

Garret, H.E. & Woodwarth, R.S. (1969). *Statistics in Psychology and Education.* Bombay : Vakils, Feffer & Simons Pvt. Ltd.

Guilford, J.P. & Benjamin, F. (1973). *Fundamental Statistics in Psychology and Education.* New York : McGraw Hill Book Co.

Srivastava, A.B.L. & Sharma, K.K. (1974). *Elementary Statistics in Psychology and Education.* New Delhi : Sterling Publishers Pvt. Ltd.