BLOCK 2
PARAMETRIC STATISTICS

ignou
THE PEOPLE'S
UNIVERSITY

UNIT 3 TEST OF SIGNIFICANCE OF DIFFERENCE BETWEEN TWO MEANS*

Structure

3.0 OBJECTIVES

After reading this unit, you will be able to:

- discuss the need and importance of the significance of the difference between means;

* Dr. Vivek Belhekar, Faculty, Department of Psychology, Bombay University, Mumbai

- explain what is null hypothesis, standard error, degrees of freedom, level of significance, two tailed and one tailed test of significance, type I error and type II error; and
- calculate the significance of difference between mean (t-test) when groups are independent, when there are correlated groups, groups matched by pair and groups matched by mean and standarddeviation.

3.1 INTRODUCTION

In psychology some times we are interested in research questions like Do the AIDS patients who are given the drug AZT show higher T-4 blood cell counts than patients who are not given that drug? Is the error rate of typist the same when work is done in a noisy environment as in a quiet one? Whether lecture method of teaching is more effective than discussion method?

Consider the question whether lecture method of teaching is more effective than discussion method. For this investigator divides the class in to two groups. One group is taught by lecture method and other by discussion method. After a few months researchers administer an achievement test for both the groups and find out the mean achievement scores of the two groups say, M_1 and M_2. The difference between these two mean is then calculated. Now the questions is whether the difference is a valid difference or it is because of sampling fluctuation or error of sampling. Whether this difference is significant or not significant. Whether on the basis of this difference, could we conclude that one method of teaching is more effective than the other method.

These questions can be answered by the statistical measures which we are going to discuss in this unit. To test the significance of difference between mean we can use either the t-test or z test. When the sample size is large, we employ z test and when sample is small, then we use the t test. In this unit we are concerned with t test. We will get acquainted with the various concepts related, to computation and description of t test.

3.2 NEED AND IMPORTANCE OF THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN MEANS

In psychology sometimes we are interested in knowing about the significance of the differences between populations. For example we are interested to discover whether ten year old boys and girls differ in their linguistic ability. Or we want to find out if children from high SES (Socio Economic Status) perform and score better academically than children from low SES. We may also try to find out at times, if two groups of persons coming from different background differ in their agility factor. Thus, many questions are asked and to be answered in psychology for which one of the measures we use is the Mean.

Let us take the first question on linguistic ability of boys and girls. First we randomly select a large sample of boys and girls (large sample means the group comprises of 30 or more than 30 persons.). Then we administer a battery of verbal test to measure the linguistic ability of the two groups and compute the mean scores on linguistic ability test of the two groups. Let us say the obtained mean scores for boys and girls are $M_1$and M_2 respectively. Now we try to find the difference between the two means. If we get a large difference ($M_1 – M_2$) in favour of the girls then we can confidently say that girls of 10 years of age are significantly more able linguistically than 10 years old boys. On the contrary if we find small difference between two means then we would conclude that ten years old girls and boys do not differ in linguistic ability.

An obtained mean is influenced by sampling fluctuation or error of sampling and whatever differences are obtained in the means, it may be due to this sampling error. Even mean of population 1 and mean of the population 2 may be the same but because of sampling error we may find the difference in the range of the two samples drawn from the two populations. In order to test the significance of an obtained difference we must first have a standard error (SE) of the difference. Then from the difference between the sample mean and standard error of difference we can say whether the difference is significant or not. Now the question arises what do we mean by significant difference? According to Garrett (1981) a difference is called significant when the probability is high and that it cannot be attributed to chance that is (temporary and accidental factors) and hence represent a true difference between population mean.

A difference is non significant when it appears reasonably certain that it could easily have arisen from sampling fluctuation and hence imply no real or true differences between the population means.

Check Your Progress I

1) When is a difference called significant?

……

……

……

……

……

3.3 FUNDAMENTAL CONCEPTS IN DETERMINING THE SIGNIFICANCE OF THE DIFFERENCE BETWEENMEANS

Let us discuss some of the fundamental concepts in determining the significance of the difference between means.

3.3.1 Null Hypothesis

This is a useful tool in testing the significance of differences. Null hypothesis asserts that there is no true difference between the two population means, and the difference found between the sample mean is therefore, accidental or unimportant (Garrett 1981). In the course of a study or an experiment, the null hypothesis is stated so that it can be tested for possible rejection. For example , to study the significant difference in linguistic ability of 8 years old girls and boys we select random sample of girls and boys and administer a battery of verbal test, compute the means of the two groups. In this study the null hypothesis may be stated thus: There exists no significant difference between the linguistic ability of boys and girls. If this null hypothesis is rejected then we can say one group is superior to the other.

3.3.2 Standard Error

The primary objective of statistical inference is to make generalisation from a sample to some population of which the sample is part. Standard error measures (1) error of sampling and (2) error of measurement. Standard error can be described as the standard deviation (SD) of all the standard deviations that have been computed for given number of samples randomly drawn from same population. Let us try to understand standard error with the help of an example. Suppose a researcher wants to carry out a research on the organisational citizenship behaviour of the employees in public sector banks. Since it in not possible to take the whole population, the researcher randomly takes around 1% of the whole population as the sample. Yet another researcher is also carrying out similar study and also draws 1% from the population. Thus, if many more researchers were carrying out such a study then they would also draw 1% sample from the whole population. The problem, though occurs when the 1% sample that is drawn by each researcher is different from each other. And though each of the 1% sample is representation of the population (that is heterogeneous in nature), they are different from each other. The researchers will also compute means and standard deviations for their respective sample. And in such a case, it is expected that the means and standard deviations would be same, because the sample has been randomly drawn from the same population. But in reality that may not happen and the mean of one sample may be lower or higher than the mean computed for another sample. The higher the standard error the less the likelihood that the sample is representative of the population. Thus, the difference between the sample needs to be close to zero so as to be sure that they represent the population.

The standard error of the mean can be calculated by the following formula:

$$SE_m \text{ or } \sigma_m = \sigma/\sqrt{N}$$

Where

σ= The standard deviation of the sample mean

N= The number of cases in the sample.

If the standard error of measurement is large it shows considerable sampling error.

3.3.3 Degrees of Freedom (df)

When a statistics is used to estimate a parameter the number of degrees of freedom (df) available depends upon the restriction placed upon the observations. One df is lost for each restriction imposed. For example we have five scores as 5,6,7,8 and 9 the mean is 7 and deviation of our scores from 7 are -2, -1, 0, 1 and 2. The sum of these deviations is zero. In consequence if any four deviations are known the remaining deviation may be automatically determined. In this way, out of the five deviations, only four (N-1) are free to vary as, the condition that the sum equals to zero impose restriction upon the independence of the 5^{th}deviation. Originally there were 5(N=5) degrees of freedom in computing the mean because all the observation or scores were independent. But as we made use of the mean for computing standard deviation we lost one degree of freedom.

Degrees of freedom varies with the nature of the population and the restriction imposed. For example in the case of value calculated between means of two independent variables, where we need to compute deviation from two means, the number of restrictions imposed goes up to two consequently df is (N- 1+N-2).

3.3.4 Level of Significance

Whether a difference between the means is to be taken as statistically significant or not depends upon the probability that the given difference could have arisen "by chance". The researcher has to take a decision about the level of significance at which he will test his hypothesis.

In social sciences .05 and .01 level of significance are most often used. When we decide to use .05 or 5% level of significance for rejecting a null hypothesis it means that the chances are 95 out of 100 that is not true and only 5 chances out of 100 the difference is a true difference.

In certain types of data, the researcher may prefer to make it more exact and use .01 or 1% level of significance. If hypothesis is rejected at this level it shows the chances are 99 out of 100 that the hypothesis is not true and only 1 chance out of 100 it is true. The level of significance which the researcher will accept should be set by researcher before collecting the data.

3.3.5 Two Tailed and One Tailed Tests of Significance

In many situations we are interested in finding the difference between obtained mean and the population mean. Our null hypothesis states that the M_1 and M_2 do not vary and the difference between them is zero. (that is,Ho:M_1- M_2=0). Whether this difference is positive or negative we are not interested in the direction of such a difference. All that we are interested is whether there is a difference. For example we hypothesised that two groups

will differ from each other we don‟t know which group will have higher mean scores and which group lower. This a non directional hypothesis and it gives rise to a two-tailed hypotheses test. In other words the difference may be in either direction and thus is said to be non directional.

In many experiments our primary concern is with the direction of the difference rather than with its existence in absolute term. For example if we are interested to determine the gain in vocabulary resulting from additions by weekly reading assignment. Here we are interested in finding out the gain in vocabulary. To take another example, if we say that training in yoga will reduce the degree of tension in persons, then we are clearly stating that there will be a reduction in the tension. In cases like this we make use of the one tailed or non directional test to test the significance of difference between the means.

3.3.6 Errors in Testing

If the null hypothesis is true and we retain it or if it is false and we reject it, we had made a correct decision. But sometimes we make errors. There are two types of errors Type I error and Type II error.

A type I error, also known as alpha error, is committed when null hypothesis (Ho) is rejected when in fact it is true. A type II error, also known as beta error, is committed when null hypothesis is retained and in fact it is not true. For example suppose that the difference between two population means (μ_1-μ_2) is actually zero, and if our test of significance when applied to the sample mean shows that the difference in population mean is significant we make a type I error. On the other hand if there is true difference between two population mean, and our test of significance based on the sample mean shows that the difference in population mean is " not significant" we commit a type II error.

Check Your Progress II

1) Given below are some statements. Indicate whether the statement is true or false:

Sr. No.	Statements	True or False
1	We commit a type I error when we reject a null hypothesis when it is really true.	
2	In testing a hypothesis, one can make three types of error.	
3	An exercise in hypothesis testing enables us to draw conclusions about the estimated parameters.	

4	For a given level of significance, we find that as the sample size increases, the critical values of t get closer to zero.	
5	If the standardised sample mean exceeds the critical clue, we should accept Ho.	

3.4 METHODS TO TEST THE SIGNIFICANCE of DIFFERENCE BETWEEN THE MEANS OF TWO INDEPENDENT GROUPS T-TEST

3.4.1 Testing Significance of Difference between Uncorrelated or Independent Means

The step used to find out significance of differences between independent mean are as below:

Step 1: Computation of the mean.

Step 2: Computation of the combined standard deviation by using the following formula:

$x_1=X_1-M_1$ (deviation of scores of first sample from its mean)

$x_2=X2-M_2$ (deviation of scores of second sample from its mean)

$$SD = \sqrt{(\Sigma x_1^2+ \Sigma x_2^2)/(N_1-1)+(N_2-1)}$$

Step 3: Computation of the standard error of the difference between two means by using following formula:

$$SE_D= SD / \sqrt{(1/N_1+1/N_2)}$$

Step 4: To Compute the t value for the difference in two independent sample mean. The following formula is used to determine t value is $t = (M_1-M_2)-0/SE_D$

Step 5: Find out the degree to freedom. The (df) degree of freedom is calculated using the following formula

$$df=(N_1-1)+(N_2-1)$$

Step 6: We then refer to table of t distributions (can be found in any Statistics book) with the calculated degree of freedom df and read the t value given under column 0.05 and 0.01of two tailed test. If our computed t value is equal or greater than the critical t value given in table then we can say that t is significant. If the computed value is lesser than given value then we will say that it is non-significant.

Let us illustrate the whole process with the help of an example

Example 1: An interest test was administered to 6 boys and 10 girls. They obtained following scores, is the mean difference between two groups significant ?

Table 3.1: Scores of boys and girls and the t value calculation

Scores obtained by boys X	x	x^2	**Scores obtained by girls (Y)**	y	y^2
28	-2	4	20	-4	16
35	5	25	16	-8	64
32	2	4	25	1	1
24	-6	36	34	10	100
36	-6	36	20	-4	16
35	5	25	28	4	16
			31	7	49
			24	0	0
			27	3	9
			15	-9	81
∑X=190		$\sum x^2$= 130	∑Y=240		$\sum y^2$ =352

Calculations

M_x= ΣX/ N = 190/ 6 = 31.66 M_y= ΣY/ N = 240/ 10 = 24

$$\mathbf{SD= \sqrt{[(\sum x^2+\sum y^2)/(N_1-1)+(N_2-1)]}}$$

$$SD = \sqrt{[(110+352)/(6-1)+(10-1)]}$$

$$= \sqrt{462/14} = \sqrt{33} = 5.74$$

$$\mathbf{SE_d = SD\ [\sqrt{(N_1+N_2)/(N_1N_2)}]}$$

$$= 5.74 \sqrt{[(16)/(60)]} = 5.74 \times .52 = 2.98$$

$$\mathbf{t = (M_1-M_2)-0/SE_D}$$

$$= (31.66-24)-0/2.98 = 2.58$$

$df=(N_1-1)+(N_2-1)$ =(6-1)+(10-1) = 14

Referring to the critical value in table for t test (that are given in the appendix of a textbook on statistics) for df= 14, we get 2.58 at the .05 and 2.58 at the.01 level, since our t is less than 2.14, therefore we will say that the mean difference between boys and girls is non significant.

Let us take another example,

Example 2: On an academic achievement test 31 girls and 42 boys obtained the following scores.

Table 3.2: Mean scores and Standard deviation

	Mean	**SD**	**N**	**df**
Boys	40.39	8.69	31	30
Girls	35.81	8.33	42	41

Is the mean difference in favour of boys and significant at the .05 level. First we will compute the pooled standard deviation by the following formula

$SD = \sqrt{[(SD_1)^2 \times (N_1-1) + (SD_2)^2 \times (N_2-1)]/(N_{1-1})+(N_{2-1})}$

SD_1= Standard Deviation of group 1 i.e. 8.69

N_1= Number of subject is group 1 i.e. 31

SD_2= Standard devotion of groups 2 i.e. 8.33

N2= Number of subject in group 2 i.e. 42

$SD = \sqrt{[(8.69)^2 \times (31-1)+(8.33)^2 \times (42-1)/(31-1)+(42-1)]} = 8.48$

$SE_d = SD\ [\sqrt{(N_1+N_2)/(N_{1x}N_2)}]$

$= 8.48\ [\sqrt{(31+42)/(31X42)}] = 2.03$

$t = (M_1-M_2)-0/SE_D$

$= (40.39-35.81)-0/2.03$

$= 2.26$

$df=(N_1-1)+(N_2-1)$

$=(31-1)+(42-1)$

$=71$

Referring to the critical value in table for t test (that are given in the appendix of a textbook on statistics) for df= 71, we find t entries of 2.00 at the 0.05 level and 2.65 at the 0.01 level. The obtained t of 2.26 is significant at .05 level but not at the 0.01 level. We may say boys academic achievement is better in comparison to that girls.

3.5 SIGNIFICANCE OF THE DIFFERENCE BETWEEN TWO CORRELATED MEANS

3.5.1 The Single Group Method

In the previous section we discussed the problem of determining the significance of difference between mean obtained by two independent groups of boys and girls.

Some time we have single group and we administer the same test twice. For example if we are intending to find out the effect of training on the students‟ educational achievement, first we take the measures of participants' educational achievement before training , then we introduce the training programme, and again we take the measure of educational achievement. In this we have single group and administer educational achievement test twice. Such type of design is known as single group design. In order to get the significance of the difference between the means obtained in the before training and after training we use the following method.

$$SE_D \text{ or } \sigma_D = [\sigma 2_{M1} + \sigma 2_{M2} - 2r\ \sigma_{M1}\ \sigma_{M2}]$$

Where

σ_{M1}= Standard error of the initials test σ_{M2}=Standard error of the finals test

r= coefficient of correlation between scores on initial test and finaltest

$t=(M_1-M_2)-0/\sigma_D$

Let us illustrate the above formula will the help of following example.

Example: At the beginning of the session an educational achievement test in maths was given to 100 IX grade students. Their mean was 55.4 and SD was 7.2. After six months an equivalent from of the same test was given and the mean was 56.9 and SD was 8.0. The correlation between scores made on the first testing and second testing was .64. Has the class made significant progress in maths during six month? We may tabulate our data in the following manner:

Table 3.3: Scores in the initial and final test of students

	Initial Test	**Final Test**
No. of students	100	100
Mean	55.4	56.9
SD	7.2.	8.0
Standard error of the mean	0.72	0.80
r12	.64	

Calculation

$$\text{SED or } \sigma_D = [\sigma^2 M1 + \sigma^2 M2 - 2r\, \sigma M1\, \sigma M2]$$

$$= [(.72)^2 + (.80)^2 - 2 \times .64 \times .72 \times .80]$$

$$= .5184 + .6400 - .7373$$

$$= .42$$

$$t = (M1-M2)-0/SED$$

$$= 1.5-0/0.42$$

$$= 3.57$$

$df=(N_1-1)$

$=(100-1)$

$=99$

Referring to the critical value in table for t test (that are given in the appendix of a textbook on statistics) for df= 99, and find that the value at 0.05 level is 1.98 and at 0.01 level is 2.63. Our obtained value is 2.35 therefore this value is significant at 0.05 level and not on 0.01 level. Here we can say that class made substantial improvement in mathematical achievement in six months.

3.5.2 Difference Method

When groups are small, then we must prefer the difference method to that given above. Let us illustrate the use of this method with the help of following example.

Example: Twelve participants were tested on an attitude scale. Then they were made to read some literature in order to change their attitude. Their attitude were again measured by the same scale. The results of the initials and final testing are as under.

Table 3.4: Results of initial and final testing of attitude

Initial condition	Final condition	Difference Cond 2- Cond 1 (Mean = 8)	X = D-M	x^2
50	62	12	4 (12-8)	16
42	40	-2	-10(-2--8)	100
35	30	-5	-13(-5-8)	169
51	61	10	2 (10-8)	4
42	52	10	2 (10-8)	4
26	35	9	-1 (9-8)	1
41	51	10	2 (10-8)	4
42	52	10	2 (10-8)	4
60	68	8	0 (8-8)	0
70	84	14	6 (14-8)	36
55	63	8	0 (8-8)	0
38	50	12	4 (12-8))	16
Total		$\sum D = 96$		$\sum x^2 = 354$

Note: The sum of scores of final condition is more than sum of initial condition therefore we subtract scores of initial condition from scores of final condition (Final condition –Initial condition) and add the score to find ΣD.

Mean = $(\sum D)/N$
= 96/12=8

$$SD_D = \sqrt{[(\sum x^2)/N-1]}$$

$$=\sqrt{[354/11]}$$

$$=5.67$$

$$SE_{MD} = SD_D/\sqrt{N}$$

$$= 5.67/\sqrt{(12)}$$

$$= 1.64$$

$$t = MD-0/SE_{MD}$$

$$=8-0/1.64$$

$$=4.88$$

df=12-1
= 11

Referring to the critical value in table for t test (that are given in the appendix of a textbook on statistics) for df = 11, we find t values of 2.20 and 3.11 at the 0.05 and at the 0.01 levels. Our t of 4.88 is far above the 0.01 level. We can conclude that participants attitude changed significantly from initial to final condition.

3.5.3 The Method of Equivalent Groups

In experiments when we want to compare the relative effect of one method of treatment over another we generally take two groups, one is known as experimental group and the other is known as control group. Here we have two groups not a single group. For the desired results these two groups need to be made equivalent. This can be done by (i) Matched pair technique or (ii) Matched groups technique. These are explained below

i) **Matched pair technique:** In this techniques matching is done by pair. Matching is done on variables which are going to affect the results of the study like age, intelligence, interest, socio-economic status.

ii) **Matched groups technique:** In this technique instead of person to person matching, matching of groups is carried out in terms of Mean and S.D.

3.5.4 Matching by Pairs

Formula for calculation of standard error of difference between mean is:

$$SE_D \text{ or } \sigma D = [(\sigma M1)^2+(\sigma M2)^2 - 2r \times \sigma M1 \times \sigma M2]$$

Here,
$\sigma M1$= Standard error of mean 1
$\sigma M2$= Standard error of mean 2

r=correlation between the two groups scores

$t = (M_1-M_2)-0 / SE_D$

Example: There are two groups X and Y of Children .72 in each group are paired child to child for age and intelligence. Both groups were given group intelligence scale and scores were obtained. After three weeks experimental group participants were praised for their performance and urged to try better. The control groups did not get the incentive. Group intelligence scale again was administered on the groups. The data obtained were as follows. Did the praise affect the performance of the group or is there a significant difference between the two groups.

Table 3.5: Results of Experimental and Control Groups.

	Experimental group	Control group
No. of children in each group	72	72
Mean score of final test	88.63	83.24
SD of final test	24.36	21.62
Standard error of the mean of final test	2.89	2.57
Correlation between experimental and control group scores	.65	

$$SE_D \text{ or } \sigma D = [(\sigma M1)^2+(\sigma M2)^2 - 2r \times \sigma M1 \times \sigma M2]$$
$$= [(2.89)^2+(2.57)^2-2x.65x\ 2.89x2.57]$$

$$t=(M_1-M_2)-O/SE_D = 6.11$$
$$t= (88.63-83.24)-0/6.11$$
$$=0.88$$

df = 72-1

= 71

Referring to the critical value in table for t test (that are given in the appendix of a textbook on statistics) for df = 71, the value at 0.05 is 2.00 and at 0.01 is 2.65. The obtained t is 0-88 therefore this value is not significant at 0.05 level and not at 0.01 level. On the basis of the results is can be said that praise did not have significant effect in stimulating the performance of children.

3.5.5 Groups Matched for Mean and Standard Deviation

When groups matched in terms of mean and SD, the following formula is used to calculate 't'.

$$SE_D = [(\sigma M1)^2 / N1] + [(\sigma M2)^2 / N2\ ]-(1-r^2\ )$$
$$t = (M1-M2)-0/\ SE_D$$

SE_D= Standard error of difference
$\sigma M1$ = Standard error of mean 1
$\sigma M2$ = Standard error of mean 2
r= Correlation between final scores of two tests

The above formula can be illustrated by the following example

Example: The 58 students of academic college and 72 students of technical college were matched for mean and SD upon general intelligence test. Then the achievement on a mechanical aptitude test was compared. The questions is do the two groups enrolled in different courses differ in mechanical ability?

Table 3.6: Results of the Academic and Technical Groups

	Academic	Technical
No. of children in each group	58	72
Mean on Intelligence GTest (Y)	102.50	102.80
SD on Intelligence Test Y	33.65	31.62
Mean on Mechanical Aptitude (X)	48.52	53.51
SD on Mechanical Aptitude X	10.60	15.36
r_2	.50	

$$SE_D = [(\sigma_{M1})^2 / N1] + [(\sigma_{M2})^2 / N2]-(1-r^2)$$

$$= [(10.60)^2/58] + [(15.35)^2/72]- (1-.25)$$

$$= 4.46$$

$$t = (M_1-M_2)-0/ SE_D$$

$$t = (53.61-48.52)-0/4.46$$

$$= 0.63$$

$df=(N_1-1)+(N_2-1)$

$=(58-1)+(72-1)$

$=128$

Referring to the critical value in table for t test (that are given in the appendix of a textbook on statistics) for df = 125(which is near to 128), the value are 1.98 at .05 level and 2.62 at 0.01 our obtained value is 0.63, which is not significant at 0.05 level. We may say that two group do not differ in mechanical aptitude.

Check Your Progress III

1) Ten persons are tested before and after the experimental procedure, their scores are given below. Test the hypothesis that there is nochange.

Before	After
60	72
52	50
61	71
36	45
45	40
52	62
70	78

51	61
80	94
65	73
72	82

3.6 LET US SUM UP

In the field of psychology some time we are interested in testing the significance of difference between two sample means.

The sample may comprise of two independent groups and single groups tested twice. Some time we have two groups matched by pair or matched for means and standard deviations. The process of determining the significance of difference between the means is different in different conditions. We may broadly summarise the procedure of calculating significance of differences between means as under.

- Establish a null hypothesis.
- Decide a suitable level of significance 0.05 or 0.01
- Determine the standard error of the difference between means of two samples.
- Compute the value of t.
- Find out the degrees of freedom.
- Determine the critical value of t from the 't' table (that is given in appendix of textbook on statistics).
- If the computed value is same or more than the value given in the table then it is taken to be significant if the computed valued is less than the given value it is considered as non significant.
- When the t-value is significant we reject the null hypothesis and when t-value is not significant we retain the null hypothesis.

3.7 REFERENCES

Kerlinger, Fred N. (1963). Foundation of Behavioural Research, (2nd Indian reprint) Surjeet Publication, New Delhi.

Garrett, H.E. (1971), Statistics in Psychology & Education, Bombay, Vakils, Seffer & Simoss Ltd.

Guilford, J.P. (1973), Fundamental Statistics in Psychology & Education, Newyork, McGraw Hill.

3.8 KEY WORDS

Null Hypothesis: A zero difference hypothesis. A statement about the status quo about a population parameter that is being tested.

Alternative Hypotheses: A hypothesis that takes a value of population parameter different from that used in the null hypothesis. It states that there is a difference in the groups on a certain characteristic that is being tested.

Type I error: An error caused by rejecting a null hypothesis when it is true.

Type II error: An error caused by failing to reject a null hypotheses when it is not true.

One tailed t test: A statistical hypothesis test in which the alternative hypothesis is specified such that only one direction of the possible distribution of values is considered. It would state there will be an increase in the performance of students after training.

Two tailed t test: A statistical hypothesis test in which the alternative hypothesis is stated in such way that it included both the higher and the lower values of a parameter than the value specified in the null hypothesis, It would state that there will be a difference (can be an increase or decrease) in the group that undergoes training.

Significance level: The probability that a given difference arises because of a chance factor or it is a true difference.

Standard error: The standard deviation of a sampling distributions.

3.9 ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress I

1) When is a difference called significant?

According to Garrett (1981) a difference is called significant when the probability is high and that it cannot be attributed to chance that is (Temporary and accidental factors) and hence represent a true difference between population mean.

Check Your Progress II

Sr. No.	Statements	True or False
1	We commit a Type I error when we reject a null hypothesis when it is really true.	True
2	In testing a hypothesis, one can make three types of error.	False
3	An exercise in hypothesis testing enables us to draw conclusions about the estimated parameters.	True

4	For a given level of significance, we find that as the sample size increases, the critical values of t get closer to zero.	True
5	If the standardised sample mean exceeds the critical clue, we should accept Ho.	False

Check Your Progress III

1) Ten persons are tested before and after the experimental procedure, their scores are given below. Test the hypothesis that there is nochange.

Before	After
60	72
52	50
61	71
36	45
45	40
52	62
70	78
51	61
80	94
65	73
72	82

t value = 4.37, significant at 0.01 level

3.10 UNIT END QUESTIONS

1) Differentiate between:

   a) Null hypothesis and alternative hypothesis.

   b) One tailed test and two tailed test.

   c) Type I error and Type II error.

2) Explain the concept of Standard error and level of significance.

3) In the first trial of a practice period, 25 twelve-year-olds have a mean score of 80.00 and a SD of 8.00 upon a digit-symbol learning test. On the tenth trial, the mean is 84.00 and the SD is 10.00. The r between scores on the first and tenth trials is .40. Our hypothesis is that practice leads to gain.

UNIT 4 TEST OF SIGNIFICANCE OF DIFFERENCE BETWEEN MORE THAN TWO MEANS*

Structure

4.0 OBJECTIVES

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

- discuss the concept of variance;
- describe the computation of one way Analysis of Variance (ANOVA);
- discuss factorial designs; and
- describe the computation of two way Analysis of Variance (ANOVA).

4.1 INTRODUCTION

A researcher wanted to study the effect of stages of adolescence, early, middle and late, on the emotional intelligence of adolescents in Chennai. For the purpose, she used the standardised psychological tests and collected data from a sample of 600 adolescents in Chennai, 200 each of early adolescents, middle adolescents and late adolescents. As the data was organised, the researcher then had to decide about which statistical technique to use for the purpose of data analysis. Initially the researcher felt that t- test can be used. But in the present case, there were three groups and as such t test is used when there are more than two groups. The research then decided to use one way ANOVA. Though it would be possible to use t test in this situation as

* Prof. Suhas Shetgovekar, Faculty, Discipline of Psychology, IGNOU, Delhi

well, but then the researcher will have to first compare early and middle adolescents, then early and late adolescents and then middle and late adolescents with regard to emotional intelligence. This will become further cumbersome if there are more than three groups, for instance ten groups and so on. Also, t test will not provide any information about the variance that may exist from the mean values of the given groups. In such a situation, one way ANOVA can be conveniently and effectively used.

In the present unit thus, we will mainly focus on the concept of variance and then we will discuss the computation of one way ANOVA and two way ANOVA. In this context we need to remember that both these techniques fall under parametric statistics and thus the assumptions of parametric statistics need to be met before these techniques are used for calculations.

4.2 CONCEPT OF VARIANCE

In BPCC104, we discussed about the concept of variance. Let us recapitulate the concept.

The term variance was used to describe the square of the standard deviation by R.A. Fisher in 1913. The concept of variance is of great importance in advanced work where it is possible to split the total into several parts, each attributable to one of the factors causing variations in their original series. Variance is a measure of the dispersion of a set of data points around their mean value. It is a mathematical expectation of the average squared deviations. It mainly helps in separating the variability in to factors that are random and factors that are systematic (Veraraghavan and Shetgovekar, 2016).

The variance (s^2) or mean square (MS) is the arithmetic mean of the squared deviations of individual scores from their means. In other words, it is the mean of the squared deviation of scores.Variance is expressed as $V = SD^2$.

The variance and the closely related standard deviation are measures that indicate how the scores are spread out in a distribution. In other words, they are measures of variability. The variance is computed as the average squared deviation of each number from its mean.

Calculating the variance is an important part of many statistical applications and analysis. It is a good absolute measure of variability and is useful in computation of Analysis of Variance (ANOVA) to find out the significance of differences between sample means.

In this context we also discuss about between group variance and within group variance. Between groups variance can be explained as variance that exists between the group means. For example, the variance in group means of early, middle and late adolescents. Whereas, within group variance can be explained as variance that exists amongst the members within a certain group. For example, variance that may exist amongst the early adolescents.

Let us now focus on some of the characteristics of variance:

1) Variance can be termed as a measure of variability and it provides information about variance in the scores in similar way as other measures of variability. You may recall that we discussed about variance in the context of measures of variability in BPCC104.

2) Variance is denoted in terms of an area. In normal probability curve, the variance is denoted in terms of area either on the right or left side of the curve. On the other hand, standard deviation is denoted by direction on the normal probability curve.

3) The value of variance is always positive.

4) The variance will remain the same, even if a certain constant in a data is subtracted or added.

The variance (s^2) or mean square (MS) is the arithmetic mean of the squared deviations of individual scores from their means. In other words, it is the mean of the squared deviation of scores.Variance is expressed as V = SD^2. The formula thus is as follows:

Variance = SD^2 or $\sigma^2 = \Sigma (X - M) / N$

Where,

X= Raw scores of a group

M= Mean of the raw scores

N= Total of the raw scores.

Variance is rigidly defined and based on all observations.It is also amenable to further algebraic treatment and is not affected by sampling fluctuations. Further, it is less erratic. The main disadvantage of variance is that it is difficult to calculate and gives greater weight to extreme values.

Check Your Progress I

1) State any one characteristic of variance.

.....

.....

.....

4.4 CONCEPT OF ANALYSIS OF VARIANCE (ANOVA)

In 1923, Ronald. A. Fisher reported about ANOVA, though the name F test was given to it by George W. Snedecor in honour of Fisher (Mohanty and Misra, 2016). As the name suggests, ANOVA is mainly related to variance rather than standard error or standard deviation. It is mainly used in order to measure the differences between two or more sample means at the same time. And it is also

how it can be termed as more advantageous when compared to t test. ANOVA can be computed for independent measures as well as repeated measures. Though the focus of the present unit will be on ANOVA for independent measures.

To just focus on the terms independent measures and repeated measures. In independent measures, for each treatment or condition there is a separate sample. The design here can be referred to as between group design. Whereas, with regard to repeated measures, the same sample is used for varied treatments or conditions.

As stated by Mohanty and Misra (2016), the meaning of analysis is division in to smaller parts and the process of carrying out the same is termed as Analysis of Variance . In this context we discuss about two relevant components, namely, between treatments variance and within treatments variance. The main focus of between treatments variance is measuring how the sample means differ from each other and this difference can either be attributed to the treatment effect or chance factor. For instance, if we say that a significant difference exits in job satisfaction of employees based on the intervention received by them namely, technique 1, technique 2 and technique 3. In this case the occupational stress may vary amongst the employees based on the intervention received by them or it could also vary due to chance factors that could be due to individual differences that may exist amongst the participants or even due to experimental error.

In ANOVA we also focus on within treatment variance. This has mainly to do with variance within a particular sample. For instance, in the example we discussed earlier where the employees were give three different techniques to measure their impact on job satisfaction of the employees. There was a sample under each treatment that received the same technique. The scores of the employees of the employees who received same technique may also vary and this can be attributed to chance factor and thus within treatment variance will help us understand to what extent a difference due to chance factor is reasonable to expect (Mohanty and Misra, 2016).

As we know, ANOVA is parametric statistical technique and thus, the assumptions of parametric statistics need to be met.

Some of the assumptions of ANOVA are as follows:

1) The F distribution of the dependent variable needs to be normal.
2) There needs to be homogeneity of variance indicating that the population from which the sample for the research is taken displays equal variance.
3) The observations need to be independent.
4) The effect of various factors on the total variations needs to be additive (not multiplicative). Thus, it can be said that in ANOVA, the observation can be divided in to independent as well as additive parts and each part can be as a result of a source that can be identified.

Check Your Progress II

1) What is between treatment variance?

.....

.....

.....

.....

4.4 COMPUTATION OF ONE WAY ANALYSIS OF VARIANCE (ANOVA)

As the concept of variance is clear, let us discuss one way ANOVA with help of an example.

A researcher was interested in studying the achievement motivation of students belonging to low, middle and high Socio Economic Status (SES). The researcher wanted to find out the effect of Socio Economic Status (SES) on self-concept of the students. In this case, the independent variable is SES with three categories or levels and the dependent variable is self-concept that is continuous in nature. Here it is not possible for the researcher to carry out t test and there are more than two categories and thus in this case one way ANOVA can be used to find out the effect of independent variable on dependent variable.

The data collected by the researcher is as follows:

Table 4.1: Example of One Way ANOVA

	Low SES (X_1)	**Middle SES (X_2)**	**High SES (X_3)**	**Total**
1	2	2	2	**6**
2	3	4	3	**10**
3	4	5	4	**13**
4	2	5	2	**9**
5	3	5	5	**13**
6	2	2	2	**6**
7	2	3	2	**7**
8	2	5	3	**10**
9	3	5	2	**10**
10	3	2	3	**8**
Total	**26**	**38**	**28**	**92**
Mean	**2.6**	**3.8**	**2.8**	

Let us us now focus on the steps involved in computation of one way ANOVA.

Step 1: Mean is to be computed for each group

As can be seen in table 4.1, the mean has been computed for each group.

For low SES, the mean is 26/ 10= 2.6

For middle SES, the mean is 38/10 = 3.8

For high SES, the mean is 28/10= 2.8

Step 2: Correction term (C) is to be computed

The formula for correction term is

$$C = (\Sigma X)^2 / N$$

Thus, in our example,

$(\Sigma X)^2$is $(26+ 38+ 28)^2 = (92)^2$

N = 30

$$C = (\Sigma X)^2 / N$$

$$= (26+ 38+ 28)^2 / 30$$

$$= (92)^2 / 30$$

$$= 8464 / 30$$

$$= 282.13$$

Thus, C is obtained as 282.13.

Step 3: Sum of Squares (SS_t) is to be computed.

Sum of Squares (SS_t) is computed by squaring each score and subtracting Correction sum.

For the purpose, we will first square each scores given in the three groups.

Table 4.2: Squaring of each score in the group

	Low SES (X_1)	**$(X_1)^2$**	**Middle SES (X_2)**	**$(X_2)^2$**	**High SES (X_3)**	**$(X_3)^2$**
1	2	**4**	2	**4**	2	**4**
2	3	**9**	4	**16**	3	**9**
3	4	**16**	5	**25**	4	**16**
4	2	**4**	5	**25**	2	**4**
5	3	**9**	5	**25**	5	**25**
6	2	**4**	2	**4**	2	**4**
7	2	**4**	3	**9**	2	**4**
8	2	**4**	5	**25**	3	**9**
9	3	**9**	5	**25**	2	**4**
10	3	**9**	2	**4**	3	**9**
Total	**26**	**72**	**38**	**162**	**28**	**88**

Let us compute Sum of Squares (SS_t) with the help of the formula

$$SS_t = [(\Sigma X_1^2 + \Sigma X_2^2 + \Sigma X_3^2) - C]$$

$$= [(72 + 162 + 88) - 282.13]$$

$$= 322 - 282.13$$

$$= 39.87$$

Thus, SS_t is obtained as 39.87.

Step 4: The between groups Sum of Squares (SS_b) is to be computed

The between groups Sum of Squares (SS_b) is computed by squaring the total of each group divided by respective number of cases, N1, N2, N3 and subtracting Correction sum (C) from the obtained value.

$$\mathbf{SS_b = (X1^2 /N1) + (X2^2 /N2) + ((X3^2 /N3) - C}$$

$$= \frac{26 \times 26}{10} + \frac{38 \times 38}{10} + \frac{28 \times 28}{10} - 282.13$$

$$= (67.6 + 14.44 + 78.4) - 282.13$$

$$= 290.4 - 282.13$$

$$= 8.27$$

Thus, SS_b is obtained as 8.27.

Step 5: Sum of Squares (SS_w) is to be computed

Sum of Squares (SS_w) is to computed by subtracting the Between sum of Square from the Total Sum of Square.

$$\mathbf{SS_w = SS_t - SS_b}$$

$$= 39.87 - 8.27$$

$$= 31.6$$

Thus, SS_w is obtained as 31.6.

Step 6: The degree of freedom (df) are worked out.

For total Sum of Squares N- 1= 30-1= 29

For between group Sum of Squares K- 1 = 3-1 = 2

For within group Sum of Squares N- K = 29- 2 = 27

Step 7: Computation of F ratio

Table 4.3: Summary of ANOVA

Source of Variance	Sum of Squares	Df	Mean Square Variance (MSS)
Between Group	8.27	2	8.27/2 = 4.14
Within group	31.6	27	31.6/27 = 1.17
Total	543.47	29	

Thus, F = MSS between groups / MSS within groups

F = 4.14/ 1.17

F= 3.54

The F ratio is obtained for the degree of freedom (df) = 2,27

Step 8: Interpretation

Now we have computed the F ratio, but then the value as such is meaningless, until and unless it is interpreted. For the purpose of interpretation, you need to refer to certain tables. These tables are given in the appendix of any Statistics books. The title of the table could be "F ratio for 0.01 and 0.05 levels of significance" (approximately) and may vary to some extent in different books. In such a table, the degree of freedom for greater mean square is given as headings for the column (on top) and degree of freedom for smaller mean square is given as headings for the rows (left hand side).

Based on the df smaller mean square and df for greater mean square, you can identify the critical value given in the table for 0.01 and 0.05 levels of significance.

In the case of our example, the df for smaller mean square as 2 and df for greater mean square as 27, the critical value is 19.50 at 0.05 level of significance and 99.50 at 0.01 level of significance. The F ratio obtained in the example 3.54 is thus not significant as the value is less than the critical values at 0.05 and 0.01 levels of significance. Thus, it can be said that there is no significance difference between low, middle and high SES with regard to achievement motivation.

Check Your Progress III

1. What is the formula for correction term?

   ……

   ……

   ……

4.5 FACTORIAL DESIGNS

In one way ANOVA, the independent variable is categorical in nature and has more than two categories or levels and the dependent variable is continuous in nature. With regard to two way ANOVA, there are two independent variables that are both categorical and can have two or more categories or levels and there is one dependent variable that is continuous in nature.

The design that is used here can be termed as factorial design. Let us discuss the same before we go on to computation of two way ANOVA.

Factorial designs are mainly used to study the effectof more than two independent variables on the dependent variable. The main effect (of each variable separately) as well as interaction effect (of all the IVs) studied with the help of this design.

Various types of factorial design are as follows:

- **2 x 2 factorial design:** Here there are two independent variables, each with two categories or levels. For example, gender (male and females) and managers (junior and senior).
- **2 x 3 factorial design:**Here there are two independent variables. One with two categories or levels and the other with three categories or levels. For example, gender (male and females) and Socio Economic Status (high, middle and low).
- **3 x 3 factorial design:**Here there are two independent variables. each with three categories or levels. For example, Socio Economic Status (high, middle and low) and stages of adolescence (early, middle and late)
- **n x k factorial design:** Here there are two independent variables with n and k categories or levels. n and k can take any number.
- **2 x 2 x 2 factorial design:** Here there are three independent variables, each with two categories or levels. For example, gender (males and females), managers (junior and senior) and Socio Economic Status (high and low).
- **3 x 3 x 3 factorial design:** Here there are three independent variables, each with three categories or levels. For example, Socio Economic Status (high, middle and low), phases of adolescents (early, middle and late) and religion (Hindus, Muslims and Christians).
- **n x k x l factorial design:** Here there are three independent variables with n and k categories or levels. n, k and l can take any number.

Two way ANOVA can help in studying the main effect and the interaction effect of the independent variables on the dependent variables. Main effect can be described as " the mean difference amongst the levels of one factor" (Mohanty and Misra, 2016, page 572). Factor means independent variable. Thus, when we display the independent variables, they would be form of columns and rows (as can be seen in table 4.4). Thus, the mean different between the rows would denote the main effect of one variable and mean difference in columns will denote the mean difference of the other variable.

Table 4.4: Hypothetical example for two way ANOVA

	Stages of adolescence (Independent Variable B)			
Gender (Independent variable A)	**Early adolescence (b_1)**	**Middle adolescence (b_2)**	**Late adolescence (b_3)**	
Male (a_1)	Mean for a_1 and b_1= 60	Mean $a_1$and b_2= 55	Mean$a_1$and b_3= 70	**Meana_1= 61.7**
Female (a_2)	Mean a_2 and b_1= 40	Mean a_2 and b_3= 65	Mean $a_2$and b_3= 80	**Meana_2= 61.7**
	Mean b_1= 50	**Mean b_2= 60**	**Meanb_3= 75**	**61.7**

As can be seen in table 4.4, the two independent variables are Stages of adolescence and gender. And in the cells we have mentioned means that would be based on the dependent variables. The mean difference between $a_1$and a_2 will constitute the main effect for the independent variable A and the mean difference between b_1 and b_2 and b_3 will constitute the main effect for independent variable A. Besides main effect, we also need to discuss about the interaction effect. This is the effect that interaction between variable A and B has on the dependent variable.

Two way ANOVA can also be effectively displayed in form of line graph. The interaction effect in this regard can be of three types, that are discussed as follows:

- **Parallel:** Parallel lines indicate that there is no interaction between the independent variables in their effect on dependent variable. For example, if we have two independent variables, gender (males and females) and SES (High and low) for dependent variable self-concept (that would be continuous in nature), then the line graph obtained could be parallel indicating that there is no interaction between the two independent variables with regard to their effect on the dependent variable that is self-concept (refer to figure 4.1). Though, when the two way ANOVA is computed, the interpretation will be based on whether the F ratio obtained is significant at 0.01 or 0.05 levels of significance or not.

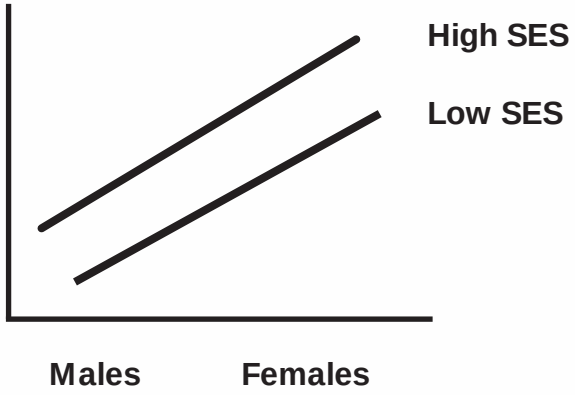

Fig 4.1: Parallel

- **Additive:** Here there is some interaction between the independent variables with regard their effect on dependent variables. In this, as you can see in figure 4.2, the lines are not parallel but as such, there is no complete interaction.

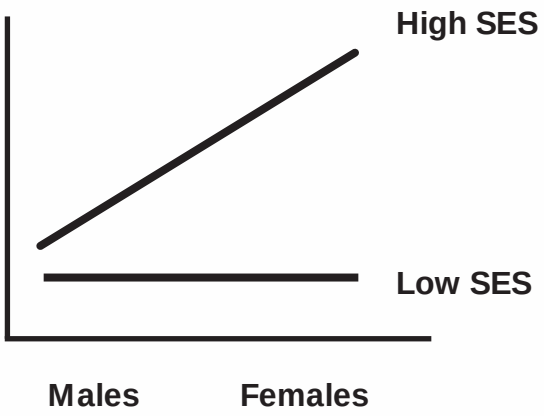

Fig 4.2: Additives

This denotes that there is some interaction between the independent variables with regard to their effect on the dependent variable. Thus, as can be seen that the self-concept of males with high SES is low when compared to females with high SES, though it is quite close to the scores obtained by males from low SES. However there may not be difference between the scores obtained by male and female high SES, but there seems to be a difference (mostly significant, that will be apparent from the F ratio) between the females of high and low SES.

- **Crossover interaction:** Here there is complete interaction between the independent variables with regard to their effect on dependent variable. As can be seen in the figure 4.3, there is a crossover interaction and the males belonging to high SES have lower self-concept as opposed to females having high SES and the males of low SES have higher self-concept compared to females of low SES.

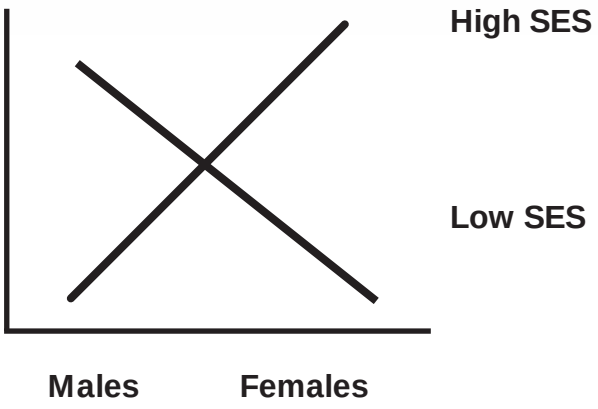

Fig 4.3: Cross over interaction

Note: The example discussed here is hypothetical and may not hold true in reality, but has been used in the context for clarity of the concept.

Check Your Progress IV

1) What is crossover interaction?

……

……

……

4.6 COMPUTATION OF TWO WAY ANALYSIS OF VARIANCE (ANOVA)

Let us us now discuss about two way ANOVA with the help of an example.

A researcher was interested in studying the perceived parental behaviour of adolescents with regard to gender, and stages of adolescence . In this case, the first independent variable is is gender with two categories or levels and the other independent variable is stages of adolescence, early, middle and late. The dependent variable is perceived parental behaviour that is continuous in nature.

Here it is not possible for the researcher to carry out a one way ANOVA as there are two independent variables that are categorical and the researcher may also be interested in finding out the interaction between the two independent variables with regard to their effect on the dependent variables. Thus in this case, the two way ANOVA can be used.

Table 4.5: Example of Two Way ANOVA-1

	Stages of Adolescence (Variable B)			**Details of rows**
Gender (Variable A)	**Early adolescence (b_1)**	**Middle adolescence (b_2)**	**Late adolescence (b_3)**	
Males (a_1)	2	2	2	
	3	4	3	
	4	5	4	
	2	5	2	
	3	5	5	
n	**5**	**5**	**5**	**n_{a1}= 15**
ΣX	**14**	**21**	**16**	**ΣX_{a1}= 51**
Mean	**2.8**	**4.2**	**3.2**	**Mean $_{a1}$= 8.07**
ΣX^2	**42**	**95**	**58**	**ΣX_{a1}^2= 195**
Females (a_2)	2	2	2	
	2	3	2	
	2	5	3	

	3	5	2	
	3	2	3	
n	**5**	**5**	**5**	**n_{a2}= 15**
ΣX	**12**	**17**	**12**	**ΣX_{a2}= 41**
Mean	**2.4**	**3.4**	**2.4**	**Mean $_{a2}$= 7.6**
ΣX^2	**30**	**67**	**30**	**ΣX_{a2}^2= 127**
Details of columns				
n	**10**	**10**	**10**	**nb_2=30**
ΣX	**26**	**38**	**28**	**ΣΣX= 92**
Mean	**2.6**	**3.8**	**2.8**	**Mean= 3.07**
ΣX^2	**72**	**162**	**88**	**$\Sigma\Sigma X^2$= 322**

As you can see, there are a total of six groups in the above example, which are as follows:

- early adolescence who are males
- middle adolescence who are males
- late adolescence who are males
- early adolescence who are females
- middle adolescence who are females
- late adolescence who are females

The sums of rows and columns will be as follows:

Gender	**Pases of Adolescents**			**Rpw Total**
	Early	**Middle**	**Late**	
Males	14	21	16	51
Females	12	17	12	41
Column total	26	38	28	92

In table 4.5, we have calculated n, ΣX, mean and ΣX^2 for both rows and columns.

Let us now try to understand the computation of ANOVA stepwise

Step 1: Compute the correction term (C) with the help of the formula

$$C = (\Sigma X)^2 / N$$

In this formula $N = n_1 + n_2 + n_3 = 10 + 10 + 10 = 30$.

Thus, N = 30

$$C = (\Sigma X)^2 / N$$

$$= (92)^2/ 30$$

$$= 8464/ 30$$

$$= 282.13$$

Thus, C is obtained as 282.13.

Step 2: Compute total Sum of Squares (SS_t)

Let us compute total Sum of Squares (SS_t) with the help of the formula

$$SS_t = \Sigma\Sigma X^2 - C$$

$$= 322\text{-}282.13$$

$$= 39.7$$

Thus, SS_t is obtained as 39.7

Step 3: Between group Sum of Squares ($SS_{between}$) is to be computed.

Between group Sum of Squares ($SS_{between}$) is to be computedwith the help of the following formula:

$$SS_{between} = \Sigma (\Sigma X)^2/n - C$$

$$= \frac{(\Sigma X_1)^2}{n_1}+\frac{(\Sigma X_2)^2}{n_2}+ \frac{(\Sigma X_3)^2}{n_3} +\frac{(\Sigma X_4)^2}{n_4}+\frac{(\Sigma X_5)^2}{n_5}+\frac{(\Sigma X_6)^2}{n_6}- C$$

$$= \frac{(14)^2}{5}+\frac{(21)^2}{5}+ \frac{(16)^2}{5}+\frac{(12)^2}{5} +\frac{(17)^2}{5} +\frac{(12)^2}{5}- 282.13$$

$$= \frac{196}{5} + \frac{441}{5}+ \frac{256}{5}+ \frac{144}{5}+ \frac{289}{5}+ \frac{144}{5} - 282.13$$

$$= 1470/5 - 282.13$$

$$= 294 - 282.13$$

$$= 11.87$$

Thus, $SS_{between}$is obtained as 11.87

Step 4:Within group Sum of Squares (SS_w) for columns is to be computed.

Within group Sum of Squares (SS_w) for columns is to be computed with the help of the following formula:

$$\mathbf{SS_w} = SS_t - SS_b$$

$$= 39.7\text{-} 12.87$$

$$= 27.43$$

Thus, $SS_w$is obtained as 27.43

Step 5: 'A' Sum of Squares (SS_a) is to be computed.

This is mainly the main effect of variable A or it can also be mentioned as the effect of the row.

'A' Sum of Squares (SS_a) is to computed

$$SS_a = \frac{(\Sigma X_{a1})^2}{n_1} + \frac{(\Sigma X_{a2})^2}{n_2} - C$$

$$SS_a = \frac{(51)^2}{15} + \frac{(41)^2}{15} - 282.13$$

$$= 2601 + 1681/15 - 282.13$$

$$= 4282/15 - 282.13$$

$$285.47 - 282.13$$

$$= 3.34$$

Thus, SS_a is obtained as 3.34.

Step 6: 'B' Sum of Squares (SS_b) is to be computed.

This is mainly the main effect of variable B or it can also be mentioned as the effect of the column.

'B' Sum of Squares (SS_b) is to computed

$$SS_b = \frac{(\Sigma X_{b1})^2}{n_1} + \frac{(\Sigma X_{b2})^2}{n_2} + \frac{(\Sigma X_{b3})^2}{n_3} - C$$

$$SS_b = \frac{(26)^2}{10} + \frac{(38)^2}{10} + \frac{(28)^2}{10} - 282.13$$

$$= 67.6 + 144.4 + 78.4 - 282.13$$

$$= 290.4 - 282.13$$

$$= 8.27$$

Thus, SS_b is obtained as 8.27.

Step 7: AB interactionSum of Squares (SSa_b) is to be computed.

$$\mathbf{(SSa_b) = SS_{between} - SS_a - SS_b}$$

$$= 11.87 - 3.34 - 8.27$$

$$= 11.87 - 11.61$$

$$= 0.26$$

Thus, SSa_b is obtained as 0.26

Step 8: Degree of freedom

In our example, variable A, that is, gender has two levels (represented by number of rows) and variable B, that is, stages of adolescence has three levels (represented by number of columns). Thus, r (rows) = 2 and c (columns) = 3. Further, there are six conditions (k) and the number of observations in each is 5 (n). The degree of free would be as follows:

$$\text{df for } SS_t = df_t = N-1 = kn-1 = 30-1 = 29$$

$$\text{df for } SS_{between} = df_{between} = k-1 = 6-1 = 5$$

$$\text{df for } SS_w = df_w = N-k = 30-6 = 24$$

The df for $SS_{between}$ can be in three parts as given below:

$$\text{df for } SS_a = df_a = r-1 = 2-1 = 1$$

$$\text{df for } SS_b = df_b = c-1 = 3-1 = 2$$

$$\text{df for } SS_{ab} = df_{ab} = (r-1)(c-1) = (2-1)(3-1) = 1x2 = 2$$

Step 9: Variance estimate or mean squares (MS) is to be computed.

Variance estimate or mean squares is to be computed with the help of following formula:

$$MS = SS/df$$

We need to compute MS for the following:

MS for variable A

$$MS_a = SS_a/df_a$$
$$= 3.34/1$$
$$= 3.34$$

MS for variable B

$$MS_b = SS_b/df_b$$
$$= 8.27/ 2$$
$$= 4.14$$

MS for AB interaction

$$MS_{ab} = SS_{ab}/df_{ab}$$
$$= 0.26/ 2$$
$$= 0.13$$

MS for within groups

$$MS_w = SS_w/df_w$$
$$= 27.43/ 24$$
$$= 1.14$$

Step 10: Computation of F ratios

F ratios are computes as follows for variable A, variable B and AB:

$$\mathbf{F_a = MS_a/MS_w}$$
$$\mathbf{= 3.34/ 1.14}$$
$$=2.93$$
$$\mathbf{F_b = MS_b/MS_w}$$
$$4.14/ 1.14$$

=3.63

$F_{ab} = MS_{ab}/MS_w$

0.13/ 1.14

=0.11

The summary of two way ANOVA is given in table 4.6

Table 4.6: Summary of Two Way ANOVA

Source of Variation	Sums of Squares SS	Degrees of freedom DF	Mean Squares MS	F	*p*-value
A	SS_a=3.34	*a*-1=1	MS_R= 3.34/1 =3.34	3.34/ 1.14 =2.93	0.1039
B	SS_b=8.27	*b*-1=2	MS_C= 8.2667/ 2 =4.14	4.14/ 1.14 =3.63	0.0449
AB	SS_{ab}=0.27	(*a*-1)(*b*-1)=2	MS_{AB}= 0.26/ 2 =0.13	0.13/ 1.14 =0.11	0.8925
Error (residual)	SS_E=28	*rab-ab*=24	MS_E= 27.43/ 24 =1.14		
Total	SS_T=39.8667	*rab*-1=29			

Step 11: Interpretation

The F ratio obtained can be interpreted based on the table of two way ANOVA that is normally given in any statistics or research methodology books.

Consulting such a table the table value for degree of freedom (df) 1, 24 , that is in the context of variable A, is 4.26 at 0.05 level of significance and 7.82 at 0.01 level of significance. The F ratio for variable A is obtained as 2.93 which is less than the table value and thus the F ratio for variable A is not

significant and null hypothesis can be accepted and it can be said that no gender difference exists with regard to perceived parental behaviour.

With regard to variable B, the F ratio is obtained as 3.63. The df here is 2, 24 and the table value is 3.40 for 0.05 level of significance and 5.61 for 0.01 level of significance. The F ratio obtained is less than the table value at 0.01 level, but more than the table value at 0.05 level. Thus, it can be said that F ratio is significant at 0.05 level of significance and it can be said that significant difference exists in perceived parental behaviour with regard to stages of adolescence .

With regard to AB interaction, the F ratio is obtained as 0.11. The df here is 2, 24 and the table value is 3.40 for 0.05 level of significance and 5.61 for 0.01 level of significance. The F ratio obtained is less than the table value at both 0.01 and 0.05 levels of significance. Thus, it can be said that F ratio is not significant.

If the AB interaction is significant, it can be shown using a line graph with mean scores on the y axis and the levels of one of the variables on the axis, with the lines of the graph showing the levels of the other variables, as shown in figures 4.1, 4.2,and 4.3.

Check Your Progress V

1) What is the formula for total Sum of Squares?

4. 7 LET US SUM UP

To sum up, in the present unit we mainly discussed about the concept of variance. The variance (s^2) or mean square (MS) is the arithmetic mean of the squared deviations of individual scores from their means. In other words, it is the mean of the squared deviation of scores. Variance is expressed as $V = SD^2$. The characteristics of variance were also discussed. Further, in the unit we explained the concept of ANOVA. In 1923, Ronald. A. Fisher reported about ANOVA, though the name F test was given to it by George W. Snedecor in honour of Fisher. The assumptions of ANOVA were also described. The unit then focused on the computation of one way ANOVA with the help of steps and example. The unit later discussed about factorial designs. Factorial designs are mainly used to study the effect of more than two independent variable(s) on the dependent variable. The main effect (of each variable separately) as well as interaction effect (of all the IVs) studied with the help of this design. Various types of factorial design were also discussed. Lastly, the unit described the computation of two way ANOVA with the help of steps and example.

4.8 REFERENCES

King, Bruce. M; Minium, Edward. W. (2008). *Statistical Reasoning in the Behavioural Sciences*. Delhi: John Wiley and Sons, Ltd.

Mangal, S. K. (2002). *Statistics in Psychology and Education*. new Delhi: Phi Learning Private Limited.

Minium, E. W., King, B. M., & Bear, G. (2001). *Statistical Reasoning in Psychology and Education*. Singapore: John-Wiley.

Mohanty, B and Misra, S. (2016). *Statistics for Behavioural and Social Sciences*. Delhi: Sage.

Veeraraghavan, V and Shetgovekar, S. (2016). *Textbook of Parametric and Non-parametric Statistics*. Delhi: Sage.

4.9 KEY WORDS

Factorial designs: Factorial designs are mainly used to study the effectof more than two independent variable (s)on the dependent variable. The main effect (of each variable separately) as well as interaction effect (of all the IVs) studied with the help of this design.

Variance: The variance (s^2) or mean square (MS) is the arithmetic mean of the squared deviations of individual scores from their means. In other words, it is the mean of the squared deviation of scores.Variance is expressed as $V = SD^2$.

4.10 ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress I

1) State any one characteristic of variance.

   Variance can be terms as a measure of variability and it provides information about variance in the scores in similar way as other measures of variability.

Check Your Progress II

1) What is between treatment variance?

   The main focus of between treatments variance is measuring how the sample means differ from each other.

Check Your Progress III

1) What is the formula for correction term?

   The formula for correction term is

$$C = (\Sigma X)^2 / N$$

Check Your Progress IV

1) What is crossover interaction?

In crossover interaction, there is complete interaction between the independent variables with regard to their effect on dependent variable.

Check Your Progress V

1) What is the formula for total Sum of Squares?

$$SS_t = \Sigma\Sigma X^2 - C$$

4.12 UNIT END QUESTIONS

1) Explain the concept of variance.
2) What is Analysis of Variance (ANOVA)
3) Describe factorial designs with suitable examples.
4) The achievement motivation scores of individuals belonging to high, middle and low Socio Economic Status are given below. Compute One way ANOVA.

High Socio Economic Status	Middle Socio Economic Status	Low Socio Economic Status
17	34	34
23	32	32
34	45	22
32	43	24
23	22	25
14	23	32
17	34	25
18	23	32
23	32	33
22	13	34

5) A researcher wanted to study the effect of gender (male and female) and area of residence (rural and urban) on self-concept of individuals. The data is as follows:

Gender	Area of residence	
	Rural	Urban
Males	2	3
	14	6
	12	7
	7	4
	6	9
Females	6	3
	6	3
	8	4
	3	7
	1	3