The *df* for "among the means of conditions" are (8 − 1) or 7, less by one than the number of conditions. The *df within* groups or within conditions are (47 − 7) or 40. This last *df* may also be found directly: since there are (6 − 1) or 5 *df* for each condition ($N = 6$ in each group), 5 × 8 (number of conditions) gives 40 *df* for within groups. The variance among *M*'s of groups is 3527/7 or 503.9; and the variance within groups is 5666/40 or 141.6.

If N = number of scores in all and k = number of categories or groups, we have for the general case that.

$$\begin{aligned} df \text{ for total } SS &= (N-1) \\ df \text{ for within groups } SS &= (N-k) \\ df \text{ for among means of groups } SS &= (k-1) \end{aligned}$$

Also: $(N-1) = (N-k) + (k-1)$

Step 6

In the present problem the null hypothesis asserts that the 8 sets of scores are in reality random samples drawn from the same normally distributed population, and that the means of conditions A, B, C, D, E, F, G and H will differ *only* through fluctuations of sampling. To test this hypothesis we divide the "among means" variance by the "within groups" variance and compare the resulting *variance ratio*, called *F*, with the *F* values in Table F. The *F* in our problem is 3.56 and the *df* are 7 for the numerator (df_1) and 40 for the denominator (df_2). Entering Table F we read from column 7 (midway between 6 and 8) and row 40 that an *F* of 2.26 is significant at the .05 level and an *F* of 3.14 is significant at the .01 level. Only the .05 and .01 points are given in the table. These entries mean that, for the given *df*'s, variance ratios or *F*'s of 2.26 and 3.14 can be expected once in 20 and once in 100 trials, respectively, when the null hypothesis is true. Since our *F* is larger than the .01 level, it would occur less than once in 100 trials by chance. We reject the null hypothesis, therefore, and conclude that the means of our 8 groups do in fact differ.

F furnishes a comprehensive or over-all test of the significance of the differences among means. A significant *F* does not tell us *which* means differ significantly, but that at least one is reliably different from some others. If *F* is not significant, there is no reason for further testing, as none of the mean differences will be significant (see p. 184). But if *F* is significant, we may proceed to test the separate differences by the *t* test (p. 191) as shown in Table 39 C.

2

The analysis of variance was developed by Sir Ronald A. Fisher, a renowned British Statistician, and the name F test was given to it by Snedecor in Fisher's honour. The variance ratios, designated as F, were tabulated by Snedecor in 1946. This device has made tremendous contribution to designing of experiments and their statistical analysis.

The analysis of variance deals with variances rather than with standard deviations and standard errors. The technique is useful in testing differences between two or more means. Its special merit lies in testing differences between all of the means at the same time. The analysis of variance is a powerful aid to the researcher. It helps him in designing studies efficiently, and enables him to take account of the interacting variables. It also aids in testing hypotheses. It provides the basis for nearly all the tests of significance in the designs which we shall consider in the following chapters. The working of the analysis of variance and its reasoning, therefore, should be thoroughly grasped. In this chapter, we will try to present some of the concepts and the working of the analysis of variance, which is the foundation of experimental design. It will help in understanding the principles of designing experiments and the statistical analysis.

ANALYSIS OF VARIANCE AND t TEST

The t test of significance is adequate when we want to determine whether or not two means differ significantly from each other. It is employed in case of experiments involving only two groups. However, for various reasons, t test is not adequate for comparisons involving more than two means. The most serious objection to the use of t, when more than two comparisons are to be made, is the large number of computations involved. For example, for three groups, $3\left(\frac{3\times 2}{2}\right)^*$ comparisons or combinations taken two at a time are required to be made and for 5 groups $10\left(\frac{5\times 4}{2}\right)$ and for 10 groups $45\left(\frac{10\times 9}{2}\right)$ comparisons are needed. Thus, as the number of groups increase the number of comparisons to be made increase rapidly, that is, the computation work increases disproportionately. Further, if a few comparisons turn out to be significant, it will be difficult to interpret the results. Let us take up an example to elucidate this point.

Suppose, in an experiment the investigator is interested in studying the effect of

*For k groups, the number of comparisons will be $\frac{k(k-1)}{2}$

10 treatments. Evidently, 45 possible t tests will have to be made for 10 treatment conditions. That is, first test $H_0 : \mu_1 = \mu_2$; then second test $H_0 : \mu_1 = \mu_3$; and so on, till we perform all the 45 t tests for the difference between every pair of means. Out of the 45 t tests, we expect to find an average of 2 or 3 t's (.05 × 45) to be significant at 5 p.c. level by chance alone. Suppose, we find that 5 differences are significant at .05 level. When t test is being applied there is no way to know whether these differences are true differences or within chance expectation. The more statistical tests we perform, for example several t tests, the more likely it is that some more differences will be statistically significant purely by chance. Thus, the t test is not an adequate procedure to simultaneously evaluate three or more means. We would like the probability of Type I error in the experiment to be .05 or less.

The analysis of variance or the F test, on the other hand, permits us to evaluate three or more means at one time. In making comparisons in experiments involving more than two means, the equality breaks down. Hence, the analysis of variance should always be preferred. The F is also an adequate test for determining the significance of two means. For two groups (df = 1), $\sqrt{F} = t$ or $F = t^2$. Therefore, in case of two treatment conditions or two groups, it is a matter of choice which one of the two tests (t or F) is used. Both yield exactly the same outcome. This means that the one-way analysis of variance and the two-tailed t test can be used interchangeably in comparing the differences between two means. However, it will be found that in the same situation F test is easier to perform than the t test.

THE CONCEPT OF VARIANCE

Variance is the very foundation of experimentation and is an extremely useful concept. Let us, therefore, try to understand its meaning and uses before handling simple analysis of variance.

Variance is a measure of the dispersion or spread of a set of scores. It describes the extent to which the scores differ from each other. The square root of the variance is called the standard deviation (s). However, because of its mathematical properties, the variance is more useful than standard deviation in research. Variance and variation, though, used synonimously are not identical terms. Variance is only one of the several statistical methods of representing variation. Variation is, thus, a more general term which includes variance as one of the methods of representing variation.

NUMERICAL EXAMPLE

The concept of variance will be explained with the help of a numerical example. Suppose, an investigator is interested in evaluating two different methods of instruction on 5th grade children. Two independent groups of 10 children each are randomly selected from a large number of children in a 5th grade class. The distribution of their achievement scores before administering the treatment (methods of instruction) is as given in Table 2.1. The scores are arranged in the ascending order.

Table 2.1 THE DISTRIBUTION OF SCORES IN THE TWO SUBGROUPS (BEFORE TREATMENT)

Subject Number	X_A	X_B
1	1	2
2	2	3
3	4	5
4	5	7
5	7	9
6	9	10
7	10	12
8	12	13
9	14	14
10	16	15
Σ	80	90

X_A and X_B represent scores of subgroups *A* and *B* respectively.

In Figure 2.1 the distribution of 10 scores in each of the two subgroups has been presented. Their means are $\bar{X}_A$ and $\bar{X}_B$ and variances s_A^2 and s_B^2 respectively.

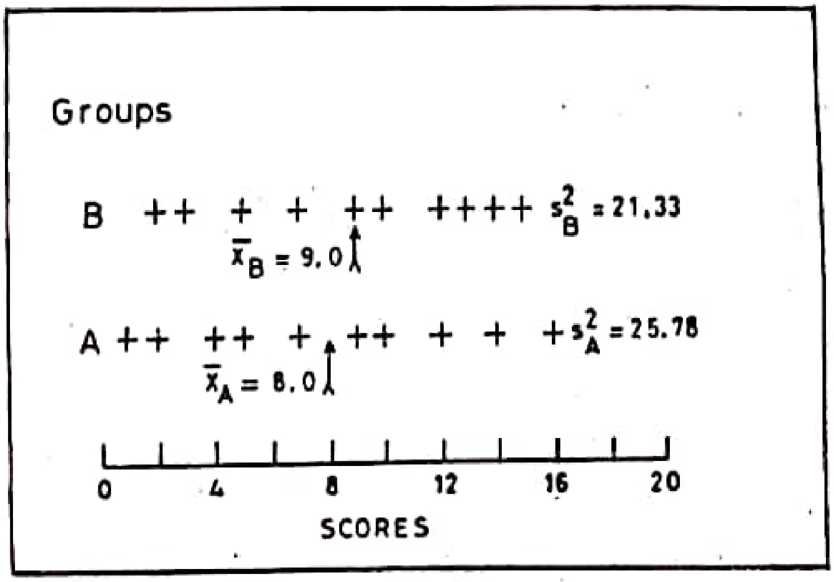

Fig. 2.1 Distribution of the scores in the two subgroups before the treatment.

Careful observation of Figure 2.1 reveals:

(i) That, the scores vary about their subgroup means. Further, the variability

($s_A^2 = 25.78$ and $s_B^2 = 21.33$) or spread of scores about their respective means is similar, within the limits of chance variation.

(ii) That, the subgroup means ($\bar{X}_A = 8.0$ and $\bar{X}_B = 9.0$) are similar, within the limits of chance variation.

The above two observations in the two samples are in accordance with the expectations of random sampling. That is, the scores in each subgroup vary about the respective means to a similar extent and, further, the subgroup means are also similar but not identical, as the two samples were selected randomly from the same population.

The investigator, then, administers the treatment to the two subgroups; treatments being assigned randomly. That is, the two subgroups are given two different methods of instruction. After a period of training, an achievement test is given to both the subgroups. The distribution of scores of the achievement test, after the application of treatment is presented in Table 2.2. The scores are arranged in the ascending order.

Table 2.2 DISTRIBUTION OF SCORES IN THE TWO SUBGROUPS (AFTER TREATMENT)

Subject Number	X_A	X_B
1	3	8
2	4	10
3	6	12
4	7	14
5	8	15
6	11	17
7	12	19
8	13	20
9	16	22
10	20	23
Σ	100	160

X_A and X_B represent scores of subgroups A and B respectively.

The distribution of scores in the two subgroups has been presented in Fig. 2.2 with their means ($\bar{X}_A$ and $\bar{X}_B$) and variances (s_A^2 and s_B^2).

The following observations can be made from Fig. 2.2, in comparison to Fig. 2.1

(i) Within each subgroup the scores vary about their subgroup means ($\bar{X}_A = 10.0$ and $\bar{X}_B = 16.0$). However, the variability in each sub-group about their respective means is not much different ($s_A^2 = 29.33$ and $s_B^2 = 25.77$), within the limits of chance variations.

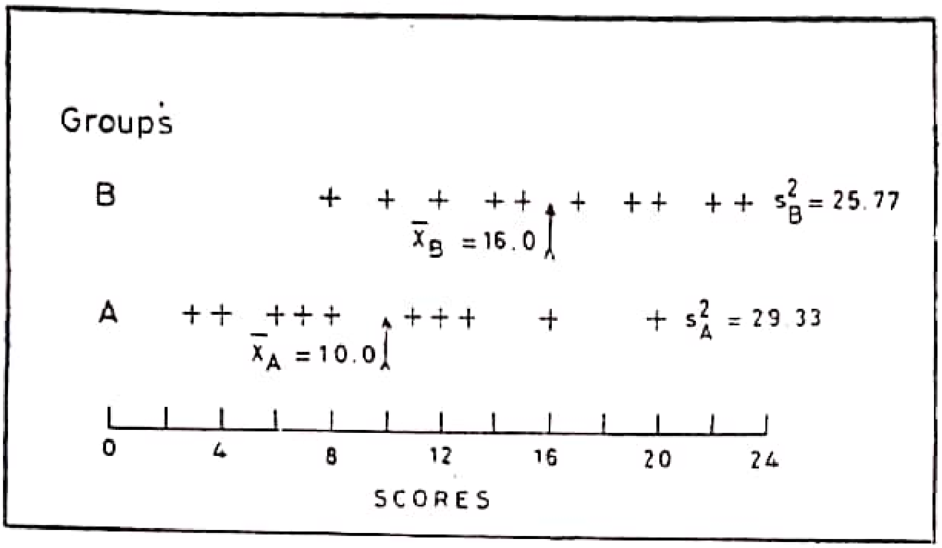

Fig. 2.2 Distribution of scores in the two subgroups after the treatment.

(ii) The subgroup means after treatment (Fig. 2.2) have drifted apart in comparison to the closeness of means of the two subgroups observed before treatment (Fig. 2.1).

The comparative analysis, before and after the treatment, in Figs. 2.1 and 2.2 respectively is particularly important for understanding the analysis of variance and the reasoning behind experimental design. Let us examine.

The variability of subgroup means is of special importance in the analysis of variance, as it reflects the variation attributable to the treatment effect as well as other uncontrolled sources of variation. Let us again refer to Fig. 2.1. We find the two means are similar, within the limits of chance variation. However, in Fig. 2.2, we can observe the effect of treatment of sub-group means; these have drifted apart. This shows that the treatment has caused variation in the subgroup means. This is called between group variation.

We have just seen that the treatment caused the subgroup means to drift apart. We have also observed (Figs. 2.1 and 2.2) that the scores within each subgroup vary about their respective means (observe the scattering of scores of the two groups around the arrow point, marking the subgroup means). This variability is also of particular importance in the analysis of variance. The pooled variability of scores about their respective subgroup means is called within group variation or "error". It is free from the influence of differential treatment.

Thus, we have been able to identify two sources of variation in the scores—one which reflects the effect of treatment is called "between group" variation and the one that reflects the variability within the subgroups is called "within groups" or "error" variation. An increase in the difference among the means results in an increase in the variance of means, and it is this variance that we evaluate relative to the error variance. The procedure adopted for this is called the analysis of variance. If the variability between the groups is considerably greater than the error variability, this is indicative of the treatment effect.

Perhaps, the most general way of classifying variability is as systematic variation and unsystematic variation. Systematic variation causes the scores to lean more in one direction than another. We observed in Fig. 2.2 that the application of treatment resulted in systematic variation in the means of the two subgroups. The variable manipulated by the experimenter is associated with systematic variation.

Unsystematic variation on the other hand, is the fluctuation in the scores due to the operation of chance and other uncontrolled sources of variation in the experiment. Random assignment of subjects in different groups helps in reducing the unsystematic variation or error.

The most important function of experimental design is to maximize the systematic variation, control the extraneous source of variation, and minimize the unsystematic or error variation. We will see in the later chapters that this objective is achieved in different ways in different designs.

A variance in the terminology of the analysis of variance is more frequently called a mean square.

$$\text{Mean Square (MS)} = \frac{\text{Variation}}{\text{df}} = \frac{\text{SS}}{\text{df}}$$

In words, a mean square is the average variation per degree of freedom. It is also the basic definition of variance.

In the foregoing discussion we have explored the concept of variance and its importance in the analysis of variance. Before we take up the computation of simple or one-way analysis of variance, let us understand some other important concepts used in the analysis of variance, like sum of squares, mean square, df, etc., and their computation.

NUMERICAL EXAMPLE

Suppose, a group of 5 subjects is given a performance test and the distribution of their scores is as given in Table 2.3.

Table 2.3 THE PERFORMANCE SCORES OF 5 SUBJECTS

Subjects	i	ii	iii
	X	$x = (X - \bar{X})$	x^2
1	2	−3	9
2	4	−1	1
3	5	0	0
4	6	1	1
5	8	3	9
	$\Sigma X = 25$ $\bar{X} = 5$	$\Sigma x = 0$	$\Sigma x^2 = 20$

$$\text{Sum of Squares or SS} = \Sigma x^2 = 20 \tag{2.1}$$

Comments

STEP 1. In column (i), the sum of the scores (ΣX) has been worked out and is equal to 25.

STEP 2. In column (ii), the scores have been squared and summed up. The sum of the squares of the scores (ΣX^2) is equal to 145.

The sum of squares (Σx^2), by the mean deviation method, was derived by adding up the square of the deviation of the scores from the group mean. However, in the above computation (direct method) we have squared the raw scores. Therefore, in order to derive the sum of squares we have to apply a correction (C). From the raw scores the sum of squares can be derived directly by applying formula 2.3. A correction term [$C=(\Sigma X)^2/N=(25)^2/5=125$] is subtracted from the sum of column (ii), i.e., ΣX^2. Thus, subtracting 125 (C) from 145 (ΣX^2), we obtain the sum of squares that is equal to 20, the value obtained by the mean deviation method also. The value of the mean square derived by both the methods is the same, which is obtained in the same manner, that is, by dividing the sum of squares by the df (SS/df), as given in formula 2.4.

It is important to understand the working of the direct method as in this book we shall always be following the direct or the raw score method. This method is preferred over the mean deviation method for its elegance and ease. This method comes handy, if a calculator is available to the investigator.

ONE-WAY ANALYSIS OF VARIANCE

We have just explored the variance notion and learnt the methods employed for computing the sum of squares and the mean square. Now, we shall try to grasp the working of the one-way analysis of variance with the help of a numerical example. In shortened form the analysis of variance is called ANOVA and sometimes ANOVAR.

The rationale of the analysis of variance is that the total variability of a set of measures, composed of several groups, can be partitioned into specific parts, each identifiable with a given source of variation. In the simple analysis of variance, the total sum of squares is partitioned into two parts: a sum of squares based upon the variation between the group means, and a sum of squares based upon variations within the several groups. On dividing the sum of squares by df, we obtain mean square, abbreviated as MS. Here the sample values are referred to as mean squares and not variance. The mean squares (sample values) are estimates of the variances (population values).

We have, thus, two estimates of the population variance—between groups and within groups. The F may, thus, be defined as

$$F = \frac{\text{Between Groups Mean Square}}{\text{Within Groups Mean Square}} \qquad (2.5)$$

The principle involved in the analysis of variance is the comparison of the variability between the various groups with the sum of the variability found within the groups. If the former variability is sufficiently larger than the latter, then it is evidence of treatment effect and we reject the null hypothesis (H_0) and accept the alternative hypothesis (H_1). However, if the difference between the sources of variability falls within the range expected from sampling error, the analysis of variance will lead to the decision of retaining the null hypothesis (H_0). We shall, then, conclude that there was no evidence of treatment effect and the differences between the group means were due to chance.

The null hypothesis which is tested by ANOVA, is that the k means of the populations from which the samples were randomly drawn are all equal, that is, $H_0 : \mu_1 = \mu_2 = \mu_3 \ldots = \mu_k$. The rejection of H_0 tells us only that some inequality exists. To investigate the inequality, we test the means pairwise. The procedure will be explained in chapter 4.

It may be noted that the decision to reject or not to reject the null hypothesis is a probabilistic decision. In the analysis of variance the decision to reject or retain the null hypothesis is made on the basis of F distribution tables, given in Appendix, Table B. In F test we need two sets of degrees of freedom (df): One for the numerator and the other for the denominator.

NUMERICAL EXAMPLE

An investigator is interested in exploring the most effective method of instruction in the class room. He decides to try three methods: Lecture (1); Seminar (2); and Discussion (3). He randomly selects 5 subjects for each of the three groups from a class of 10th grade students. After three months of instructions an achievement test is administered to the three groups. The distribution of achievement scores in the three groups is as given in Table 2.5.

Table 2.5 THE DISTRIBUTION OF ACHIEVEMENT SCORES OF SUBJECTS TREATED BY THE THREE METHODS OF INSTRUCTIONS

Subject Number	Method: Lecture (1)	Seminar (2)	Discussion (3)	
1	8	11	5	
2	10	13	5	
3	11	13	8	
4	11	15	9	
5	12	16	10	
Σ	52	68	37	157 G

Here $n = 5$; $k = 3$; $N = kn = 5 \times 3 = 15$

Partitioning of Total Variation and df

In the simple analysis of variance, the total variation and df will have the following partitioning:

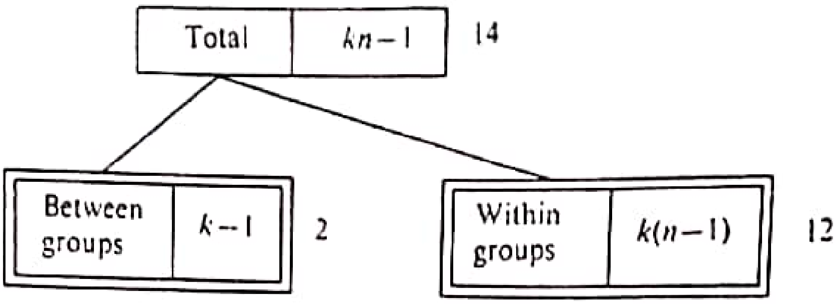

Fig. 2.3 Schematic representation of the analysis.

where

n = number of subjects in each of the three subgroups ($n = 5$)

k = number of subgroups ($k = 3$)

kn = total number of subjects or observations in the experiment (N)

The left hand rectangles indicate the partitioning of the sum of squares and the adjoining rectangles indicate the partitioning of total df, in the general form. The numerals outside the rectangles are the df associated with the numerical example. The double-line enclosed rectangles indicate the final partitioning.

In the equation form the partitioning of the total sum of squares may be expressed as

$$SS_{total} = SS_{bet.groups} + SS_{w.groups} \tag{2.6}$$

where

SS_{total} = total sum of squares generated from the deviation of the individual observations from the mean of the total observations in the experiment

$SS_{bet.groups}$ = between groups sum of squares generated from the deviation of the subgroup means from the mean of the total observations in the experiment

$SS_{w.groups}$ = within groups sum of squares generated from the pooled deviation of the individual observation from the respective subgroup means

Computation

(i) Correction Term $(C) = \frac{G^2}{kn} = \frac{(157)^2}{15} = 1643.27$

(ii) Total SS $= (\Sigma X^2) - C$

$= (8^2 + 10^2 + 11^2 + \ldots + 9^2 + 10^2) - C$

$= 1785.00 - 1643.27 = 141.73$

(iii) Between Groups SS $= \frac{\Sigma(\Sigma X)^2}{n} - C = \frac{52^2}{5} + \frac{68^2}{5} + \frac{37^2}{5} - C$

$= 1739.4 - 1643.27 = 96.13$

(iv) Within Groups SS = Total SS − Between Groups SS

$= 141.73 - 96.13 = 45.6$

(v) Table 2.6 SUMMARY OF ONE-WAY ANALYSIS OF VARIANCE

Source of Variation	SS	df	MS	F
Between Groups	96.13	2	48.07	12.65**
Within Groups (Error)	45.60	12	3.80	
Total	141.73	14		

$^{**}F_{.01}(2, 12) = 6.93$ $\quad F = \frac{48.07}{3.80} = 12.65$

Comments

Partitioning of Total Variation and df

In the one-way analysis of variance the total sum of squares is partitioned into two component parts—one due to the variation between the groups and the other due to the variation within the groups. In the present problem the total degrees of freedom are 14 ($kn - 1$), partitioned into two component parts, 2 df [$(k - 1)$] attributable to the variation between the groups and 12 df [$k(n - 1)$] to the variation within the groups.

An important aspect of the analysis of variance is the partitioning of total sum of squares and degrees of freedom. It will be observed later that the partitioning differs with the nature of the design. Once, the partitioning of sum of squares and df of a design is understood, the computation part is mechanical. Therefore, before starting the actual analysis work, one should try to comprehend the schematic representation of the analysis, and follow the computations step by step.

Computation

STEP 1: *Correction Term* As explained earlier, for computing the sum of squares by the direct method a correction is needed. The correction term (C) was, however, the same for deriving all the sum of squares in the numerical example, with the exception of within groups sum of squares, explained under the comments in Step 4.

The correction term is obtained by squaring the grand total ($G = 52 + 68 + 37 = 157$) and then dividing it by the total number of subjects or observations in the experiment ($N = kn = 3 \times 5 = 15$). The correction term (G^2/N), was found to be equal to 1643.27.

STEP 2: *The Total SS* The total sum of squares is a measure of the total variation of the individual scores about the combined mean. It reflects all the sources of variation, that is, between groups variation and within groups variation in the present case.

The total sum of squares is obtained by combining the scores of the three groups and treating them as one set of scores. In Step 2 each of the 15 raw scores is first squared, then the squares are summed, and thereafter correction term is subtracted from the obtained sum. The total sum of squares in the present example is equal to 141.73.

In Step 2, all scores have not been displayed; the ommission of certain terms of the sequence has been indicated by dots. For example, $8^2 + 10^2 + \ldots + 9^2 + 10^2$ indicates that the individual scores from the first to the last of the distribution have been squared and added.

STEP 3: *Between Groups SS* The sum of squares between groups is a measure of the variation of the group means about the combined mean. If the group means do not differ among themselves at all the sum of squares between groups will be zero. Thus, greater the variation in the group means, the larger is the sum of squares between groups.

In Step 3, between groups sum of squares has been obtained by the direct method. The totals of each of the three subgroups (i.e., 52, 68, and 37) have been squared and divided by the number of observations in each subgroup and summed $[\Sigma(\Sigma X)^2/n]$. Finally, the correction term (C) has been subtracted from the sum of squares. The between groups sum of squares is found to be equal to 96.13.

STEP 4: *Within Groups SS* The within group sum of squares is the pooled sum of squares based on the variation within each group about its own mean. The within groups sum of squares is also called error sum of squares. All the uncontrolled sources of variation are pooled in the within groups sum of squares.

In Step 4, the sum of squares within groups has been obtained by subtraction, taking advantage of the addition theorem characterizing this analysis. From equation 2.6 it is observed that

$$SS_{total} = SS_{bet.\ groups} + SS_{w.\ groups}$$

$$\therefore \quad SS_{w.\ groups} = SS_{total} - SS_{bet.\ groups}$$

By substituting the obtained values of SS_{total} or the total sum of squares and $SS_{between}$ or the sum of squares between groups, we obtain sum of squares within groups. It is equal to 45.6. However, there can be no verification of the computation of the within groups sum of squares by the subtraction method. Therefore, beginners would do well to calculate independently the within groups sum of squares. Let us carefully observe the computation of the within groups sum of squares by the direct method.

We have just learnt that the within groups sum of squares is the pooled sum of squares based on the variation of the individual observations about the mean of the particular subgroup. Therefore, the sum of squares within groups is equal to

$$\text{SS within subgroup 1} = (8^2 + 10^2 + \ldots + 12^2) - \frac{52^2}{5}$$

$$= 550.0 - 540.8 = 9.2$$

Note: The lower case letter n represents the number of observations or subjects in the subgroup ($n = 5$), upper case letter N represents the total number of observations or subjects in the subgroups ($N = 15$), and k represents the number of groups or treatments ($k = 3$).

$$\text{SS within subgroup 2} = (11^2 + 13^2 + \ldots + 16^2) - \frac{68^2}{5}$$

$$= 940.0 - 924.8 = 15.2$$

$$\text{SS within subgroup 3} = (5^2 + 5^2 + \ldots + 10^2) - \frac{37^2}{5}$$

$$= 295 - 273.8 = 21.2$$

$$\therefore \quad SS_{within} = 9.2 + 15.2 + 21.2 = 45.6$$

If no mistake is committed, the outcome by the direct method is exactly the same as obtained by the subtraction method.

Note: (a) The correction factor for each subgroup is different, that is, the square of the respective subgroup total divided by n, the number of observations in each subgroup.

(b) $SS_{w.\,group}$ is the sum of the individual sum of squares within each of the three subgroups.

(c) The df associated with the sum of squares within groups is also the pooled df of the subgroups, that is, 4 + 4 + 4 = 12.

STEP 5: *Analysis of Variance Summary Table* Preparing analysis of variance table is the final step in the analysis. Note carefully the format of the table 2.6 which is of standard form. First column is for sources of variation, then SS, followed by df, MS, and finally F.

The reader will recall that the mean square (MS) or variance estimate is obtained by dividing the SS by the appropriate df. Dividing the SS by its df gives an estimate of the common population variance that is indepedent of the variation of the group. Thus, dividing the between groups sum of squares by its df, i.e., 96.13 by 2 gives MS which in this example is equal to 48.07. This value is the estimate of the common population variance independent of the variation within groups. Similarly, dividing of the within groups SS by its df, i.e., 45.6 by 12 gives MS which is found to be 3.8. Again, this value is the estimate of the common population variance which is independent of the variation in the group means.

Then, the ratio (F) of the MS between groups and the MS within groups is obtained by dividing 48.07 by 3.80. Here the obtained value of F is equal to 12.65. It is entered in the first row under column F.

Test of Significance

The next step is to evaluate the obtained F value. We consult the F table in the Appendix, Table B, for 2 and 12 degrees of freedom. First we move along the top row, where degrees of freedom for greater mean square are given, and pause at 2. Then, we proceed downwards in column 2 until we find the row entry corresponding to df 12. The values of F significant at the 5 p.c. point are given in light face type, and those significant at 1 p.c. in bold (dark) face type. The critical value of F corresponding to 2 and 12 df at $\alpha = .01$ is 6.93. Since our obtained value of F, 12.65 far exceeds the critical or tabled value, 6.93, we reject the null hypothesis (H_0). The overall F indicates that the means of the three groups do not fall on a straight line with zero slope. Hence the null hypothesis that the three groups are random samples from a common normal population is rejected. On the basis of the results of the experiment, we can conclude that the three methods of instruction produced significant differences in the three groups. As F is an overall index, further tests on means have to be carried to compare the pairs of means. This aspect will be discussed in Chapter 4.

STRENGTH OF ASSOCIATION

The significant F indicates that the observed differences between the treatment means are not likely to arise by chance. However, it does not indicate anything about the strength of the treatment effect. The statistic omega square (ω^2) is a measure of the strength of treatment effect. It gives us the proportion of the total variability in a set of scores that can be accounted for by the treatments. That is, what portion of the variance in the scores can be accounted for by the differences in the treatment groups. The formula for strength of association is

$$\omega^2 = \frac{SS_{between} - (k-1)MS_{within}}{SS_{total} + MS_{within}} \tag{2.7}$$

Let us now compute the strength of treatment effects in our numerical example. The values of $SS_{between}$, SS_{total}, and MS_{within}, have been obtained from Table 2.6. The steps in computing ω^2 are given below

$$SS_{between} = 96.13$$

$$SS_{total} = 141.73$$

$$MS_{within} = 3.8$$

$$k = 3 \text{ (treatments)}$$

$$\therefore \quad \omega^2 = \frac{96.13 - (3-1)(3.8)}{141.73 + 3.8} = .608$$

Thus, approximately 61 per cent of the variance in the dependent variable is accounted for by the difference in the method of instruction. In other words, there is fairly strong relationship between methods of instruction and achievement scores of the subjects.

GENERAL COMMENTS

One may wonder, why do we keep the between groups variance in the position of the numerator and the within groups variance in the denominator. The logic is simple. If the group means are significantly different, then the mean square between groups should be larger than the mean square within groups (error). It is rare that small values of $F (F < 1)$ indicate anything but sampling variation. It is only large values of F that suggest treatment effects. Therefore, we refer to the F table only when the ratio is greater than one. If the mean square between groups is smaller than the mean square within groups, then the F value will be less than one. In the analysis of variance summary table we simply ignore the value of obtained F and there is no need to refer to the F tables as the data offer no evidence against the null hypothesis.

The significant F indicates that the three methods of instruction did produce differences in the achievement scores of the groups. However, F does not indicate which of the three differences among the pair of means are significant. To find this, post hoc comparisons between the subgroup means is done. There are a variety of methods for comparing the individual means. Some of these will be discussed in chapter 4.

SUMMARY OF STEPS

We have just completed the computation of one-way or simple analysis of variance with detailed comments on the various steps involved. Let us summarize the steps involved

1. Observe carefully the partitioning of the total variation (SS) and df.
2. Calculate the correction term (C).
3. Calculate the total sum of squares (SS_{total}).
4. Calculate the between groups sum of squares ($SS_{between}$).
5. Calculate the within groups sum of squares (SS_{within}) by subtraction or by direct method.
6. Determine the between groups df.
7. Determine the within groups df by subtraction or directly.
8. Prepare analysis of variance summary table.
9. Enter the obtained values of $SS_{between}$ and SS_{within} and their respective df.
10. Compute the between groups mean square ($MS_{between}$) by dividing $SS_{between}$ by its df.
11. Compute the within groups mean square (MS_{within}) by dividing SS_{within} by its df.
12. Calculate F ratio ($MS_{between}/MS_{within}$).
13. Compare the obtained F ratio with the critical F value from the F table.

After presenting a detailed discussion and computation of one-way analysis of variance, we now proceed a step forward to understand the rationale and computation of two-way analysis of variance.