ANALYSIS OF VARIANCE (ANOVA)

Anthony S. Morfe_CAS_Math Cluster

OBJECTIVE:

- At the end of this presentation, students should be able to use One-Way ANOVA technique to determine if there is a significant difference among three or more means.

ONE-WAY ANALYSIS OF VARIANCE (ANOVA)

- When an F test is used to test a hypothesis concerning the means of three or more populations, the technique is called analysis of variance (commonly abbreviated as ANOVA).
- H_o: $\mu_1 = \mu_2 = \mu_3 = \ldots = \mu_n$

  H_1: At least one of the means is different from the others

Note: No need to identify if left, right or two tailed test should be used since ANOVA works on more than two groups.

HOW TO USE THE F-DISTRIBUTION TABLE TO FIND THE CRITICAL VALUE

Find the critical value if d.f.N. =6 (degrees of freedom of the numerator), d.f.D.=5 (degrees of freedom of the denominator) and α=0.01.

Table H *(continued)*

	$\alpha = 0.01$											
d.f.D.: degrees of freedom, denominator	**d.f.N.: degrees of freedom, numerator**											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**12**	**15**
1	4052	4999.5	5403	5625	5764	585[illegible]	5928	5982	6022	6056	6106	6157
2	98.50	99.00	99.17	99.25	99.30	9[illegible].33	99.36	99.37	99.39	99.40	99.42	99.43
3	34.12	30.82	29.46	28.71	28.24	2[illegible].91	27.67	27.49	27.35	27.23	27.05	26.87
4	21.20	18.00	16.69	15.98	15.52	1[illegible].21	14.98	14.80	14.66	14.55	14.37	14.20
5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	10.67	10.46	10.29	10.16	10.05	9.89	9.72
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31

Answer: C.V.=10.67

Anthony S. Morfe_CAS_Math Cluster

ASSUMPTIONS FOR THE F TEST FOR COMPARING THREE OR MORE MEANS

1. The populations from which the samples were obtained must be normally or approximately normally distributed.
2. The samples must be independent of one another.
3. The variances of the populations must be equal.

FINDING THE *F* TEST VALUE FOR THE ANALYSIS OF VARIANCE

Step 1 Find the mean and variance of each sample.

$$(\bar{X}_1, s_1^2), (\bar{X}_2, s_2^2), \ldots, (\bar{X}_k, s_k^2)$$

Step 2 Find the grand mean.

$$\bar{X}_{\text{GM}} = \frac{\Sigma X}{N}$$

Step 3 Find the between-group variance.

$$s_B^2 = \frac{\Sigma n_i(\bar{X}_i - \bar{X}_{\text{GM}})^2}{k - 1}$$

where: $\bar{x}_i$=sample mean per group
s_i^2=sample variance per group
n_i=sample size per group
N=sum of the sample sizes of the groups
k=number of groups

Anthony S. Morfe_CAS_Math Cluster

FINDING THE *F* TEST VALUE FOR THE ANALYSIS OF VARIANCE

Step 4 Find the within-group variance.

$$s_W^2 = \frac{\Sigma(n_i - 1)s_i^2}{\Sigma(n_i - 1)}$$

Step 5 Find the F test value.

$$F = \frac{s_B^2}{s_W^2}$$

The degrees of freedom are

d.f.N. $= k - 1$

where k is the number of groups, and

d.f.D. $= N - k$

where N is the sum of the sample sizes of the groups

$$N = n_1 + n_2 + \cdots + n_k$$

ONE-WAY ANALYSIS OF VARIANCE (ANOVA)-EXAMPLE 1

A researcher wishes to try three different techniques to lower the blood pressure of individuals diagnosed with high blood pressure. The subjects are randomly assigned to three groups; the first group takes medication, the second group exercises, and the third group follows a special diet. After four weeks, the reduction in each person's blood pressure is recorded. At $\alpha = 0.05$, test the claim that there is no difference among the means. The data are shown.

ONE-WAY ANALYSIS OF VARIANCE (ANOVA)-EXAMPLE 1(CONT'D)

Medication	Exercise	Diet
10	6	5
12	8	9
9	3	12
15	0	8
13	2	4

ONE-WAY ANALYSIS OF VARIANCE (ANOVA)-EXAMPLE 1(CONT'D)

Medication	Exercise	Diet
10	6	5
12	8	9
9	3	12
15	0	8
13	2	4
$\overline{X}_1 = 11.8$	$\overline{X}_2 = 3.8$	$\overline{X}_3 = 7.6$
$s_1^2 = 5.7$	$s_2^2 = 10.2$	$s_3^2 = 10.3$
$n_1=5$	$n_2=5$	$n_3=5$

sample mean ($\overline{x}_i$) and sample variance (s_i^2) are computed using the shortcut in the calculator

ONE-WAY ANALYSIS OF VARIANCE (ANOVA)-EXAMPLE 1(CONT'D)

Step 1: State the hypotheses and identify the claim.

H_0: $\mu_1 = \mu_2 = \mu_3$ (claim)

H_1: At least one mean is different from the others.

Step 2: Find the critical value. Since k=3 (number of groups) and N=15 (total number of data values),

d.f.N. = k-1 = 3-1 = 2 d.f.D. = N-k = 15-3 = 12

The critical value is 3.89 obtained from the f-distribution table with $\alpha = 0.05$.

Step 3: Compute for the test value.

a. Find the mean and variance of each sample (these values are shown on the previous slide)

ONE-WAY ANALYSIS OF VARIANCE (ANOVA)-EXAMPLE 1(CONT'D)

b. Find the grand mean. The *grand mean,* denoted by $\bar{X}_{GM}$, is the mean of all values in the samples.

$$\bar{X}_{GM} = \frac{\Sigma X}{N} = \frac{10 + 12 + 9 + \cdots + 4}{15} = \frac{116}{15}$$

When samples are equal in size, find $\bar{X}_{GM}$ by summing the $\bar{X}$'s and dividing by k, where k = the number of groups.

c. Find the between-group variance, denoted by s_B^2.

$$s_B^2 = \frac{\Sigma n_i(\bar{X}_i - \bar{X}_{GM})^2}{k-1}$$

the numerator is also called sum of squares between groups (SS_B)

$$= \frac{5(11.8 - \frac{116}{15})^2 + 5(3.8 - \frac{116}{15})^2 + 5(7.6 - \frac{116}{15})^2}{3-1} = 160.1333333 / 2$$

$$= 80.06666667$$

Note: This formula finds the variance among the means by using the sample sizes as weights and considers the differences in the means.

ONE-WAY ANALYSIS OF VARIANCE (ANOVA)-EXAMPLE 1(CONT'D)

d. Find the within-group variance, denoted by s_W^2.

$$s_W^2 = \frac{\Sigma(n_i - 1)s_i^2}{\Sigma(n_i - 1)}$$

the numerator is also called sum of squares within groups (SS_W)

$$= \frac{(5-1)(5.7) + (5-1)(10.2) + (5-1)(10.3)}{(5-1) + (5-1) + (5-1)} = 104.8 / 12$$

$$= 8.733333333$$

Note: This formula finds an overall variance by calculating a weighted average of the individual variances. It does not involve using differences of the means.

e. Find the *F* test value.

$$F = \frac{s_B^2}{s_W^2} = 80.06666667/8.733333333 = 9.17$$

Note: Reject H_o if test value ≥ critical value while Do not reject H_o if test value < critical value.

Step 4 Make the decision. The decision is to reject the null hypothesis, since T.V. > C.V. $9.17 > 3.89$.

Step 5 Summarize the results. There is enough evidence to reject the claim and conclude that at least one mean is different from the others.

ONE-WAY ANALYSIS OF VARIANCE (ANOVA)

Table 12–1 Analysis of Variance Summary Table

Source	Sum of squares	d.f.	Mean square	F
Between	SS_B	$k-1$	MS_B	
Within (error)	SS_W	$N-k$	MS_W	
Total				

In the table,

SS_B = sum of squares between groups → Numerator of the between group variance (s_B^2)

SS_W = sum of squares within groups → Numerator of the within group variance (s_W^2)

k = number of groups

$N = n_1 + n_2 + \cdots + n_k$ = sum of sample sizes for groups

$$MS_B = \frac{SS_B}{k-1}$$

$$MS_W = \frac{SS_W}{N-k}$$

$$F = \frac{MS_B}{MS_W}$$

ONE-WAY ANALYSIS OF VARIANCE (ANOVA)-EXAMPLE 1(CONT'D)

Table 12–2 Analysis of Variance Summary Table for Example 1

Source	Sum of squares	d.f.	Mean square	F
Between	160.13	2	80.07	9.17
Within (error)	104.80	12	8.73	
Total	264.93	14		

REFERENCE:

Bluman, A. (2012). *Elementary Statistics: A Step by Step Approach*,8e. McGraw-Hill Higher Ed.

MCL Malayan Colleges Laguna A MAPÚA SCHOOL

MALAYAN COLLEGES LAGUNA ONLINE BETTER OUTCOMES

MITL

End of Lecture 11

Anthony S. Morfe_CAS_Math Cluster