Chapter 12

ANalysis **O**f **VA**riance

Lecture 1

Sections: 12.1 – 12.2

ANOVA test is an extension of two sample independent t-test demonstrated how to compare differences of means between two groups. The t-test is a useful tool for comparing the means of two groups; however, with three or more groups, the t-test is not an effective statistical tool. On a practical level, using the t-test to compare many means is a cumbersome process in terms of the calculations involved. On a statistical level, using the t-test to compare multiple means can lead to biased results.

There are many kinds of questions in which we might want to compare the means of several different groups at once. For example, in evaluating the effects of a particular social program, we might want to compare the mean outcomes of several different program sites. Or we might be interested in examining the relative performance of different members of a corporate sales team in terms of their monthly or annual sales records. Alternatively, in an organization with several different sales managers, we might ask whether some sales managers get more out of their sales staff than others.

With questions such as these, the preferred statistical tool is ANOVA. There are some similarities between the t-test and ANOVA. Like the t-test, ANOVA is used to test hypotheses about differences in the average values of some outcome between two groups; however, while the t-test can be used to compare two means or one mean against a known distribution, ANOVA can be used to examine differences among the means of several different groups at once. More generally, ANOVA is a statistical technique for assessing how nominal independent variables influence a continuous dependent variable.

We will focus on one-way ANOVA, a statistical tool that is used to compare multiple groups of observations, all of which are independent but may have a different mean for each group. A test of importance for many kinds of questions is whether or not all the averages of a set of groups are equal.

Definition:

A ***treatment** or **factor*** is a property, or characteristic, that allows us to distinguish the different populations from one another.

Assumptions

1. The populations have distributions that are approximately normal.
2. The populations have the same variance σ^2 or standard deviation σ.
3. The samples are simple random samples. That is, samples of the same size have the same probability of being selected.
4. The samples are independent of each other. Not matched or paired.
5. The different samples are from populations that are categorized in only one way. This is the basis for the name of the method: *one-way* analysis of variance.

Hypotheses

H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \ldots$
H_1: At least one mean is different.

Test Statistic for One-way ANOVA:

$$F = \frac{\text{Variance between samples}}{\text{Variance within samples}} = \frac{MS(Treatment)}{MS(Error)}$$

Caution when interpreting results:

When we conclude that there is sufficient evidence to reject the claim of equal population means, we cannot conclude from ANOVA that any particular mean is different from the others.

MS Treatment

Variance between samples also called variation due to treatment.

It is an estimate of the common population variance σ^2 that is based on the variation among the sample *means.* It tells us how much the means of the different groups differ from one another or we can look at it as how the individual group means differ from the overall mean.

MS Error

Variance within samples also called **variation due to error.**

It is an estimate of the common population variance σ^2 based on the sample *variances.* It gives us an estimate of how much the individual scores in a group differ from one another or differ from the mean of the group. This is the pooled variance s_p^2

Degrees of Freedom:

***k* = number of samples, *n* = sample size and**
***N* = total number of values in all samples combined**
Numerator degrees of freedom = $k - 1$
Otherwise known as df treatment

When working with equal sample sizes:
Denominator degrees of freedom = $k(n - 1) = N - k$
Otherwise known as df errror

When working with unequal sample sizes:
Denominator degrees of freedom = $N - k$
Otherwise known as df error

ANOVA Table:

Source	df	SS	MS	F	P
Treatment	k – 1	$SS_{treatmet}$	$MS_{treatment}$	MS_{tr}/MS_{er}	
Error	N – k	SS_{error}	MS_{error} s_p^2		
Total	N – 1	SS_{total}			

$$MS(treatment) = \frac{SS(treatment)}{k-1} \quad MS(error) = \frac{SS(error)}{N-k}$$

Note:

1. $N - 1 = (k - 1) + (N - k)$

2. $SS_{treatment} + SS_{error} = SS_{total}$

SS(total), or total sum of squares, is a measure of the total variation in all of the sample data combined.

SS(treatment), also referred to as SS(factor) or SS(between groups), is a measure of the variation *between* the sample means.

SS(error), also referred to as SS(within groups) or SS(within samples), is a sum of squares representing the variation that is assumed to be common to all the populations being considered.

MS(treatment) is a mean square for treatment.

MS(error) is a mean square for error. s_p^2

1. Four groups of five students each were taught a skill by four different teaching techniques. At the end of a specified time the students were tested and their scores were recorded. Do the following data indicate that there is a significant difference in the mean achievement for the four teaching techniques? Use $\alpha = 0.025$

Group-A	Group-B	Group-C	Group-D
64	73	61	63
73	82	79	69
69	71	71	68
75	69	73	74
78	74	66	75

2. An independent random sample of households in the four U.S. regions yielded the data on last year's energy consumptions shown in the table below. At a 0.05 level of significance, do the data provide sufficient evidence to conclude that a difference exists in last year's mean energy consumption by households among the four U.S. regions?

Northeast	Midwest	South	West
15	17	11	10
10	12	7	12
13	18	9	8
14	13	13	7
13	15		9
	12		

3. A district sales manager would like to determine whether there are significant differences in the mean sales among the four franchise stores in her district. Sales (in thousands of dollars) were tracked over 5 days at each of the four stores. Use $\alpha = 0.01$

Store 1	Store 2	Store 3	Store 4
10	20	3	30
15	20	7	25
10	25	5	30
20	15	10	35
20	20	4	30

4. Three brands of cigarettes, six from each brand, were tested for tar content. After some research, a pooled estimate for the standard deviation is known to be 3.743, and the total variation of all sample data is known to be 352.28. Fill out the table below.

Source	df	SS	MS	F	P
Treatment					
Error					
Total					