Sourav Verma
Research Scholar, OPJS
University, Churu, Rajasthan,
India

Role of analysis of variance (One way-anova) in music

Sourav Verma

Abstract
Testing a New Technique but not sure how it stack up against all the alternative is very difficult as most of the options sound similar to each other so picking the best out of them is a challenge. Consider a scenario we have three music treatment to be apply on student to see their reaction as music enhance the concentration and consequently help student absorb new information I have to found kind of music would be good choice for this. I conduct a survey on randomly selected student from three different classes. Each Classroom was provided with a different environment for student to study in order to make a confident and reliable decision we need a evidence to support our approach. I used ANOVA technique for making best decision. ANOVA is defined as a statistical measure that analyze the variance of two, three or more comparable series for determining the significance of the difference in their mean values with the help of statistical technique known as F-test. The methodology was originally developed by Sir Ronald A. Fisher, the pioneered innovator of the use and application of statistical methods in experiment design who coined the name "Analysis of variance"-ANOVA.

Keywords: MSC, MSE, F-test, degree of freedom

1. Introduction
The most important and fundamental application of F-distribution is the Analysis of Variance or ANOVA. It is defined as a statistical measure that analyze the variance of two, three or more comparable series for determining the significance of the differences in their mean values with the help of statistical technique known as F-test.

In others words, technique of ANOVA is used to test the equality of three or more sample means by comparing the sample variance using F-distribution. This technique split up the variance into two parts, namely
i)   Variance between the samples
ii) Variance within samples

2. Characteristics of anova
The characteristics of the analysis of variance (ANOVA) are as follows:
i)   It makes statistical analysis of variance of two, three or more samples.
ii) It determines whether the difference in the mean values of the different samples is due to
iii) Chance, or due to any significant cause and thereby it reveals the true characteristics of the given series.
iv) It gives the desired result by finding the appropriate variance ratio through the F-test technique.

3. Assumptions of analysis of variance
The analysis of variance (ANOVA) is made on the basis of some assumptions which are as follows:
i)   Each sample is drawn from the normally distributed population. If samples are large, this assumption is not required.
ii) The samples drawn at random for study are homogeneous and independent of each other.
iii) (iii. There is no significant difference amongst the variance of the different universe from which the samples have been drawn.
iv) It begins with the null hypothesis that $\mu_1 = \mu_2 = \mu_3 = \ldots \ldots = \mu_k$.

Corresponding Author:
Sourav Verma
Research Scholar, OPJS
University, Churu, Rajasthan,
India

v) It is assumed that critical values of the variance ratio, (F) as estimated by the Snedecar for

Different degrees of freedom at different level of significance viz.,5% or 1% etc. holds good Testing the significance of the calculated values of F-that represents the variance ratio of the
Sample.

4. Technique of analysis of variance (anova)
For the sake of clarity this technique be divided into two catagories, namely.
i) One-way-classification
ii) two-way- classification

But in this paper i am using only one-way classification.

4.1 One –way classification
In one–way classification, the given data is classified according one criterion (attribute) only. consider the null hypothesis.
H_0: $\mu_1 = \mu_2 = \mu_3 = \ldots = \mu_k$ i.e., means of all samples are equal and
Alternative hypothesis:
H_0: $\mu_1 \neq \mu_2 \neq \mu_3 \neq \ldots \neq \mu_k$ i.e., all means are not equal.

Working Rule for Analysis of variance
1. **Find variance between the samples:** the variance between samples measures the difference between the sample mean of each group and overall mean weighted by the number of observations in each group. It is denoted by SSC and also known as sum of squares between samples (or between the columns).

The steps is calculating the variance between the samples are as follows:
i) Calculate the mean of each sample i.e $\overline{x}_1, \overline{x}_2 \ldots \overline{x}_k$ etc.
ii) Calculate the grand mean $\overline{x}$. Its value is obtained as follows:
$\overline{x} = (\overline{x}_1 + \overline{x}_2 + \overline{x}_3 + \ldots \overline{x}_k)/k$, where k is the number of samples
iii) Take the difference between the grand mean and the means of the various samples.
iv) Square these deviations and multiply these squares by the size of the respective sample and then find the total which will gives sum of squares between the samples i.e, SSC.

Divide the total obtained in step (iv)by the degree of freedom v_1. The degree of freedom v_1 is one less than the no of samples i.e., $v_1 = k -1$, where k = the number of samples.

2. **Find the variance within the samples**: The variance within samples measures those inter-sample differences due to chance only. It is denoted by the SSE and also known as sum of squares within samples (or sum of squares due to errors).

The steps in calculating variance within samples are as follows:
i) Calculate the mean value of each samples i.e $\overline{x}_1, \overline{x}_2 \ldots \overline{x}_k$ etc.
ii) Take the difference of the various items in a sample from the mean values of the respective Samples.

iii) Square these derivations and find the total which gives the sum of square within the samples i.e., SSE
iv) Divide the total obtained in step (iii) by the degree of freedom V_2. The degree of freedom V_2 is obtained by subtracting the number of samples from the total number of items i.e, $v_2 = N-k$, where k=the number of samples and N=the total number of all the observations.

3. **Find the F-ratio and compare the calculated value of F with the table value of F**

$$F = \frac{\text{Mean sum of squares between samples (MSC)}}{\text{Mean sum of squares within samples (MSE)}}$$

Compare this calculated value of F with the table of F for the degrees of freedom at some critical level (take 5% level of significance, if not mentioned). If the calculated value of F is greater than the table value, then it is concluded that the difference in sample means is significant i.e., it could not have arisen due to fluctuations of simple sampling or the samples are not drawn from the sample population, and if the calculated value of F is less than the table value, the difference is not significant and has arisen due to fluctuations of simple sampling.

5. Learning objectives
a) Be able to identify when analysis of variance (ANOVA) should be used and what the null and alternative (research) hypotheses are.
b) Be able to identify the factor(s) and response variable from a description of an experiment.
c) Understand the basic logic of an ANOVA. Be able to describe between-sample variability (measured by mean square for groups (*MSG*)) and within-sample variability (measured by mean square error (*MSE*)).
d) Know how to compute the *F* statistic and determine its degrees of freedom given the following summary statistics: sample sizes, sample means and sample standard deviations. Be able to use technology to compute the *p*-value for *F*.
e) Be able to use technology to produce an ANOVA table.
f) Recognize that statistically significant differences among population means depend on the size of the differences among the sample means, the amount of variation within the samples, and the sample sizes.
g) Recognize when underlying assumptions for ANOVA are reasonably met so that it is appropriate to run an ANOVA.
h) Be able to create appropriate graphic displays to support conclusions drawn from ANOVA output.

6. Problem statement
While performing one-way ANOVA I will now consider the "Variable Music" treatment for simplicity of conclusion. Rather factor age I applied this test on students of different age groups predict that students will learn better if they study to Classical music (constant sound) compared to mixed music (pop, traditional, jazz Variable sound) and in third one there is no music (no sound) I randomly divide 27 students into three groups and send them to three different rooms. In the first room, students hear classical music, in the second room they hear mixed music and in the third one there is no music. Each group is given 30 minutes to study a section of their text after which they take a 10-question multiple-choice exam. Table contains the results.

Table 1.

Classical music	Mixed Music	No Music
8	5	4
5	7	7
5	3	6
7	5	2
8	5	4
7	8	3
8	4	5
3	5	5
10	7	4

6.1 Anova on excel summary

Table 2.

Groups	Count	Sum	Average	Variance
Classical music	9	61	6.777777778	4.444444444
Mixed music	9	49	5.444444444	2.527777778
No music	9	40	4.444444444	2.277777778

Table 3.

Source of variation	SS	Df	Ms	F
Between Groups	24.6666	2	12.3333333	
witin Groups	74	24	3.083333	4
Total	98.666666	26		

Steps to perform one way anova in excel

1. Input data into column or rows in excel for example if three groups of student for music treatment are being tested, spread the data into three column
2. Click the "data "tab and then click "data analysis "
3. Click "ANOVA single Factor "and then click "OK"
4. Type an input range into the input range box
5. Select an output range
6. Choose an alpha level. For most hypothesis tests, .05 is standard
7. Click "OK" The result of ANOVA will appear in a worksheet

Line Chart showing Relationship of Average and Variance of all the three groups

6.2 Manual solution using one way anova

Table 4.

©Classical Music	Mixed Music	No Music
8	5	4
5	7	7
5	3	6
7	5	2
8	5	4
7	8	3
8	4	5
3	5	5
10	7	4
Total 61	49	40

(A) Variance between the sample

$\bar{x}_1 = 6.78$
$\bar{x}_2 = 5.44$
$\bar{x}_3 = 4.44$
$\bar{\bar{x}} = (\bar{x}_1 + \bar{x}_2 + \bar{x}_3)/k$

$= (6.78 + 5.44 + 4.44) / 3$
$= (16.66)/3$
$= 5.553$

Sum of squares between samples

SSC $= \sum n_1(\bar{x}_1 - \bar{\bar{x}})^2 + \sum n_2(\bar{x}_2 - \bar{\bar{x}})^2 + \sum n_3(\bar{x}_3 - \bar{\bar{x}})^2$
$= 9(6.78\text{-}5.553)^2 + 9(5.44\text{-}5.553)^2 + 9(4.44\text{-}5.553)^2$
$= 24.73$

Mean sum of square between the samples

$$MSC = \frac{SSC}{k-1} = \frac{24.73}{3-1} = 12.363$$

(B) Variance within the samples

Table 1.

A		B		C	
x_1	$(x_1 - \bar{x}_1)^2$	x_2	$(x_2 - \bar{x}_2)^2$	x_3	$(x_3 - \bar{x}_3)^2$
8	1	5	0	4	0
5	4	7	4	7	9
5	4	3	4	6	4
7	0	5	0	2	4
8	1	5	0	4	0
7	0	8	9	3	1
8	1	4	1	5	1
3	16	5	0	5	1
10	9	7	4	4	0
Sum	36	**Sum**	22	**Sum**	20

Total Sum of square within the samples
SSE = 36 + 22 + 20 = 78

Mean sum of square within the samples
$MSE = \frac{SSE}{n-k} = \frac{78}{27-3} = 3.25$

F=MSC/MSE=12.363/3.25=3.804

Tabulated value of f of degree for freedom (2, 24) at 5% level of significance is 3.24. Calculated value 3.804 > 3.24

7. Result of the experiment
1. We have to decide whether these three group performed differently because of different situation or by merely chance
2. In a statistical sense how different these three samples from each other

Above calculation shows that the mean score of the student listening classical music is better than the other two groups, so the treatment must be helpful.
The table value of F at 5% level of significance with degree of freedom 2 and 24 is 3.40 and the calculated value of F is 3.84(appox 4) >3.40 i.e, the calculated value of F.>the table value at 5% level with degree of freedom 2 and 24.
Hence we Reject the null hypothesis H0 at 5% level of significance and conclude that the population mean are not equal at least one of the three samples have significantly different mean and thus belong to an entirely different population.

8. Key Terms
A factor is a variable that can be used to differentiate population groups. The levels of a factor are the possible

values or settings that a factor can assume. The variable of interest is the response variable, which may be related to one or more factors.

An analysis of variance or ANOVA is a method of inference used to test whether or not three or more population means are equal. In a one-way ANOVA there is one factor that is thought to be related to the response variable.

An analysis of variances tests the equality of means by comparing two types of variation, between-groups variation and within-groups variation. Between-groups variation deals with the spread of the group sample means about the grand mean, the mean of all the observations. It is measured by the mean square for groups, *MSC* Within-groups variation deals with the spread of individual data values within a group about the group mean. It is measured by the mean square error, *MSE*.

The *F*-statistic is the ratio *MSC*/MSE.

9. References

1. A Treatise on music of Hindostan: Capt. Willard. Lexicon Universal encyclopedia: lexicon Publication, New York, 1983.
2. Furnham A, Strbac L. Music is as distracting as noise: The differential distraction of background music and noise on the cognitive test performance of introverts and extraverts. Ergonomics. 2002; 45(3):203-17.
3. Aczel AD. Complete Business Statistics, Irwin, 1989.
4. Brown M, Forsythe A. Robust tests for the equality of variances. Journal of the American Statistical Association. 1974, 364-367.
5. Ostertagová E. Applied Statistic (in Slovak), Elfa, Košice, 2011.
6. Rafter JA, Abell ML, Braselton JP. Multiple Comparison Methods for Means, SIAM Review. 2002; 44(2):259-278.
7. Rykov VV, Balakrishnan N, Nikulin MS. Mathematical and Statistical Models and Methods in Reliability, Springer, 2010.
8. Stephens LJ. Advanced Statistics demystified, McGraw-Hill, 2004.