Conception on One-Way ANOVA Technique with the assist of Illustration

Amit Rakshit
Assistant Professor, Department of Mechanical Engineering, Kanad Institute of Engineering and Management, Mankar, Purba Bardhaman

Abstract- **The article demonstrates the overview of One-Way ANONA using the aid of illustration. One-Way ANOVA is shown using by direct method and short-cut method for testing the difference among different groups of samples and to investigate the most expedient method among them. In many works of ANOVA technique the stands are as follows: "The core of ANOVA is that the total amount of variation in a set of data is broken down into two types, that amount which can be attributed to chance and that amount which can be attributed to particular causes".**

Index Terms- **One-Way ANOVA, Direct method, Short-Cut method.**

1. INTRODUCTION

Analysis of variance i.e. ANOVA is one of the most vital technique used for multi criteria decision, also known as multi-objective decision making concerning in the field of researches such as engineering, education, economics, management, industry, agricultural and several other disciplines. This technique is useful when multiple sample cases are occupied. In the earlier, the means of difference between two samples can be identified with the help of either z-test or t-test, but the complexity can be arise when we happen to evaluate the significance of the difference amongst more than two sample means at the same time. The ANOVA technique enables us to perform this simultaneous test and as such is considered to be an important tool of analysis in the hands of a researcher.

2. Definition of ANOVA

Professor R.A. Fisher, who was the first man used the term 'Variance' and give details momentarily the theory of ANOVA in the realistic field. After that Professor Snedecor and many other respective dignitaries contributed to the enlargement of this technique. The name "Analysis of Variance" is more representative a procedure for testing the difference among different groups of data for homogeneity.

Analysis of Variance is a general statistical tool that is used to analyze a wide range of research designs and to investigate many multifarious problems. There may be variation between samples and within samples. ANOVA consists in splitting the variance for analytical purposes. Hence, it is a method of analysing the variance to which a response is subject to various components corresponding to various source of variation.

3. ANOVA Technique

ANOVA technique can be designed by the following customs which are as follows:

- One-way (single factor) ANOVA
- Two-way (double factor) ANOVA
- Latin-Square ANOVA.

4. One-way ANOVA

In the article we are discussing about one-way ANOVA with the help of applications. Using the one-way, we consider only one factor and then observe that the reason for thought factor to be important is that several possible types of samples can occur within the factor and we determine the difference within the factor if it is happen. The one way ANOVA is also called a single factor analysis of variance because there is only one independent factor constraints. The one way ANOVA can be fundamentally designed by direct method and short cut method.

4.1 Direct Method for One-Way ANOVA

Mathematical steps of direct method are as follows:

Step.1. Firstly, we have to obtain the mean of each sample i.e., obtain $\bar{X}_1$, $\bar{X}_2$, $\bar{X}_3$, $\bar{X}_4$,.........., $\bar{X}_K$ where there are K samples.

Step.2. Then find out the mean of the sample means as follows:

$\bar{\bar{X}}$= ($\bar{X}_1$+$\bar{X}_1$+$\bar{X}_1$+$\bar{X}_1$+............+$\bar{X}_1$)/K;

where K=number of samples.

Step.3. Take the deviations of the sample means from the mean of the sample means and calculate the square of such deviations which may be multiplied by the number of items in the respective sample and then obtain their total. This is called as the sum of squares for variance between the samples which is also denoted by SS between. Mathematically represented by:

SS between= $n_1(\bar{X}_1 - \bar{\bar{X}})^2 + n_2(\bar{X}_2 - \bar{\bar{X}})^2 + n_3(\bar{X}_3 - \bar{\bar{X}})^2 + \ldots\ldots\ldots\ldots + n_k(\bar{X}_k - \bar{\bar{X}})^2$

Step.4. Divide the result of SS between with the help of step (3) by the degree of freedom (DOF) between the samples to obtain mean square (MS) or variance between samples. Mathematically represented by:

MS between= $\left(\frac{SS\ between}{K-1}\right)$; where (k-1) represents degree of freedom (DOF) between samples.

Step.5. Take the deviations of the values of the sample items for all the samples from corresponding means of the samples and calculate the squares of such deviations and then obtain their total. This is called as sum of squares for variance within samples which is also denoted by SS within. Mathematically represented by:

SS within= $\sum(X_{1i} - \bar{X}_1)^2 + \sum(X_{2i} - \bar{X}_2)^2 + \sum(X_{3i} - \bar{X}_3)^2 + \ldots\ldots\ldots\ldots + \sum(X_{ki} - \bar{X}_k)^2$, i= 1,2,3,4,........

Step.6. Divide the result of SS within with the help of step (5) by the degree of freedom (DOF) within the samples to obtain mean square(MS) or variance within samples. Mathematically represented by:

MS within = $\left(\frac{SS\ within}{K-1}\right)$; where (n-k) represents degree of freedom (DOF) within samples,

N = total number of items in all the samples i.e., $n_1 + n_2 + n_3 + n_4 + \ldots\ldots\ldots + n_k$;

K= number of samples.

Step.7. Now, the sum of squares for total variance (SS for total variance) should be equal to the total of the result step(3) and step(5) which is explained above, i.e.,

SS for total variance= SS between + SS within.

Step.8. Finally, F- ratio may be worked out as under:

F-ratio = $\frac{MS\ between}{MS\ within}$

4.2 Procedure of table for one-way ANOVA using direct method

The computational table or decision matrix for one-way ANOVA may be worked out as under:

Source of variation	Sum of squares	Degree of freedom	Mean square	F-ratio
Between samples	$n_1(\bar{X}_1 - \bar{\bar{X}})^2 + n_2(\bar{X}_2 - \bar{\bar{X}})^2 + n_3(\bar{X}_3 - \bar{\bar{X}})^2 + \ldots\ldots\ldots\ldots + n_k(\bar{X}_k - \bar{\bar{X}})^2$	(k-1)	$\frac{SS\ between}{k-1}$	$\frac{MS\ between}{MS\ within}$
Within samples	$\sum(X_{1i} - \bar{X}_1)^2 + \sum(X_{2i} - \bar{X}_2)^2 + \sum(X_{3i} - \bar{X}_3)^2 + \ldots\ldots\ldots\ldots + \sum(X_{ki} - \bar{X}_k)^2$; i= 1,2,3,4,........	(n-k)	$\frac{SS\ within}{n-k}$	
Total	SS between + SS within	(n-1)	-	-

Illustration: Illustration of using One-Way ANOVA

Below are given the yields per acre of wheat for six plots entering a crop competition, there of the plots being sown with wheat of variety A and variety B. The table shown below:

Yields in fields per acre

Variety	I	II	III
A	30	32	22
B	20	18	16

Set up a table of analysis of variance and calculate F-ratio.

Computations:

The problem can be solved by using direct method for one-way ANOVA. Let us solve the problem with the help of given table which is mentioned below:

Table No: I Yields in fields per acre

Variety	I	II	III	Total

A	30	32	22	84
B	20	18	16	54
Total	50	50	38	138

In the above table, we can observe that the yields per acre in the variety of both A and B should be the total of same. Now, we can calculate the mean of each of these samples:

$\bar{X}_1 = \frac{(30+20)}{2} = 25$,

$\bar{X}_2 = \frac{(32+18)}{2} = 25$, and,

$\bar{X}_3 = \frac{(22+16)}{2} = 19$

Mean of the sample means

$\bar{\bar{X}} = \overline{(}\bar{X}_1 + \bar{X}_2 + \bar{X}_3)/k$; where, k= number of samples

$= (25+25+19)/3$

$= 23$

Now we find out SS between and SS within samples:

SS between $= n_1(\bar{X}_1 - \bar{\bar{X}})^2 + n_2(\bar{X}_2 - \bar{\bar{X}})^2 + n_3(\bar{X}_3 - \bar{\bar{X}})^2 + ... + n_k(\bar{X}_k - \bar{\bar{X}})^2$

$= \{2(25-23)^2 + 2(25-23)^2 + 2(19-23)^2\}$

$= (8 + 8 + 32)$

$= 48$

SS within $= \sum(X_{1i} - \bar{X}_1)^2 + \sum(X_{2i} - \bar{X}_2)^2 + \sum(X_{3i} - \bar{X}_3)^2 + ... + \sum(X_{ki} - \bar{X}_k)^2$;

where i= 1,2

$= \{(30-25)^2 + (20-25)^2\} + \{(32-25)^2 + (18-25)^2\} + \{(22-19)^2 + (16-19)^2\}$

$= (25 + 25 + 49 + 49 + 9 + 9)$

$= 166$

SS total variance can be worked out as under:

SS total = SS between + SS within

= 48 + 166

= 214

It can now set up the table of ANOVA for this particular illustration:

Table No: II

Source of variation	Sum of Squares(SS)	Degree of Freedom(DOF)	Mean Square(MS)	F-ratio
Between Samples Within Samples	48 166	(3-1) =2 (6-3)= 3	$\frac{48}{2}$= 24 $\frac{166}{3}$= 55.33	$\frac{24}{55.33}$ = 0.433
Total	214	5	79.33	-

4.3 Short Cut Method for One-Way ANOVA

In the short cut method, ANOVA can be performed when means of the sample means happen to be non-integer values. The method can be more expedient as compare to the direct method for one way ANOVA.

Mathematical steps of short cut method are as follows:

Step.1. Firstly, obtain the total values of all individual items in all the samples i.e., $T=\sum X_{ij}$; where i=1,2,3,4,………… and j=1,2,3,4,……….

Step.2. Find out the correction factor which is as under:

Correction factor = $(T^2)/n$

Step.3. Find out the square of all the sample values separately and then find its total. Deduce the correction factor from this total value and the result obtained is the sum of squares (SS) for total variance.

Mathematically represented by:

Total SS = $\sum X_{ij}^2 - (T^2)/n$; where i=1,2,3,4,......... and j=1,2,3,4,……….

Step.4. Now, obtain the square of each sample of total $(T_j)^2$ and divide the value of each sample by the number of items in the concerning sample and take the result thus obtained. Deduce the correction factor from this result which is the sum of squares (SS) between variance of the samples.

Mathematically represented by:

SS between = $\sum(T_j)^2/n_j - (T)^2/n$; where j=1,2,3,4,…………

Step.5. Finally, the sum of squares (SS) within the samples can be obtained by deduction the result of sum of squares (SS) between the samples from the total variance of the sum of squares (SS) i.e.,

SS within $= \{\sum X_{ij}^2 - (T^2)/n\} - \{\sum(T_j)^2/n_j - (T)^2/n\}$

$= \{\sum X_{ij}^2 - \sum(T_j)^2/n_j\}$

4.4 Procedure of table for one-way ANOVA using short-cut method

The computational table or decision matrix for one-way ANOVA may be worked out as under:

Source of variation	Sum of squares	Degree of freedom	Mean square	F-ratio
Between samples	$\sum(T_j)^2/n_j - (T)^2/n$	(k-1)	$\frac{SS\ between}{k-1}$	$\frac{MS\ between}{MS\ within}$

Within samples	$\{\sum X_{ij}^2 - \sum(T_j)^2/n_j\}$	(n-k)	$\frac{SS\ within}{n-k}$	
Total	$\sum X_{ij}^2$ - $(T^2)/n$	(n-1)	-	-

Illustration: Illustration of using One-Way ANOVA
Below are given the yields per acre of wheat for six plots entering a crop competition, there of the plots being sown with wheat of variety A and variety B. The table shown below:

Yields in fields per acre

Variety	I	II	III
A	30	32	22
B	20	18	16

Set up a table of analysis of variance and calculate F-ratio.

Computations:
The problem can be solved using by short cut method for one-way ANOVA. Let us solve the problem with the help of given table which is mentioned below:

Table No: III Yields in fields per acre

Variety	I	II	III	Total
A	30	32	22	84
B	20	18	16	54
Total	50	50	38	138

In the above table, we can observe that the yields per acre in the variety of both A and B should be the total of same. Now, we can calculate through short cut method:
In the above case, we first find out the total of all individual values of n samples and find out T.
Here, Total (T) in the given case = 138 and n= 6
Hence, the correction factor = $(T)^2/n$
= $(138)^2/6$
= 3174
Now, find out the total of SS and SS between which is worked out as under:
Total SS = $\sum X_{ij}^2 - (T^2)/n$; where i=1,2 and j= 1,2,3
= $\{(30)^2 + (20)^2 + (32)^2 + (18)^2 + (22)^2 + (16)^2\}$ – (3174)
= (900 + 400 + 1024 + 324 + 484 + 256) – (3174)
= (3388 – 3174)
= 214
SS between = $\sum(T_j)^2/n_j - (T)^2/n$; where j=1,2,3
= $\{(50)^2/2 + (50)^2/2 + (38)^2/2\}$ – (3174)
= (1250 + 1250 + 722) – (3174)
= (3222 –3174)
= 48
Therefore,
SS within = $\{\sum X_{ij}^2 - \sum(T_j)^2/n_j\}$
= (3388-3222)
= 166.
Note: It may be seen that exactly same result as we had obtained in the case of direct method. From now onwards we can set up ANOVA table or decision matrix and interpret F- ratio in the same way as we have done in the direct method.
It can now set up the ANOVA table for this particular illustration:

Table No: IV

Source of variation	Sum of Squares(SS)	Degree of Freedom(DOF)	Mean Square(MS)	F-ratio
Between Samples Within Samples	48 166	(3-1) =2 (6-3)= 3	$\frac{48}{2}$= 24 $\frac{166}{3}$= 55.33	$\frac{24}{55.33}$ = 0.433
Total	214	5	79.33	-

5.CONCLUSION

In this article, it can be observed that the accurately same results can be produced under direct method and short cut method. To find illustrations under the methods that satisfactorily shows its overview logic and applications. Here it can be observed that the short cut method for one way ANOVA is more expedient rather than direct method for one way ANOVA. So, it is effectively method to the activities of research domain.

REFERENCES

[1] Rohan K. Gavade, "Multi Criteria Decision Making: An overview of different selection problems and methods", International Journal of

Computer Science and Information Technology, Vol.5(4), 2014; pp. 5643-5646.

[2] M. Kaladhar, K.Venkata Subbaiah and Ch.Srinivasa Rao; "Determination of Optimum Process Parameters during turning of AISI 304 Austenitic Stainless Steels using Taguchi method and ANOVA", International Journal of Lean Thinking, Vol. 3, Issue.1, 2012; pp. 1-19.

[3] Zlatko Pavic and Vedran Novoselac; " Notes on TOPSIS method", International Journal of Research in Engineering and Science, Vol.1, Issue.2, 2013; pp. 5-12.

[4] G.H.Tzeng and J.J.Huang; "Multi Attribute Decision Making: Methods and Applications", CRC Press, NewYork, 2011.

[5] C.R. Kothari; "Research Methodology: Methods and Applications"; New Age International Publishers (formerly Wiley Eastern Limited).

[6] S.Kalavathy; "Operations Research with C Programs"; Vikas Publishing House Private Limited.