IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 6 Issue 1, January 2019
ISSN (Online) 2348 – 7968
www.ijiset.com

Two Way Analysis of Variance: The Fundamental Concepts with the aid of Applications

Amit Rakshit [1], Sumana Ghosh [2], Supriyo Chatterjee [3]

[1] Assistant Professor; Department of Mechanical Engineering;
Kanad Institute of Engineering and Management, Mankar, Purba Bardhaman, W.B
[2] Lecturer; Department of Civil Engineering;
Kanad Institute of Engineering, Mankar, Purba Bardhaman, W.B
[3] Assistant Professor; Department of Mechanical Engineering;
Seacom Engineering College, Sankrail, Howrah, W.B

Abstract:
The article demonstrates two-way ANOVA using the assist of applications. The scheme of conducting two-way ANOVA which can be the statistical tool using for an analysis of various complex phenomena in the various engineering disciplines. Here in the application of two way ANOVA we had to study the process of marketing flotation regarding repeated values and non repeated values and work out their variation. However, this choice of ANOVA technique was dictated by the adjustment of methodology to the nature of the analyzed phenomena.

Keywords: ANOVA; Two Way ANOVA; ANOVA table; Repeated values; Non repeated values

1. Introduction:

Inconsistency is a property which characterizes both controlled and uncontrolled factors influencing each stage of the experiment and its final result. A number of random factors i.e. aberrations appearing during the experiment could disturb the value of measurements. As a result more complicated phenomena is being analyzed, the bigger is the probability of measuring diversity of some features. The problem was considered by renounced Professor Ronald A.Fisher who was the first man to use the term "Variance" and, in fact, it was he who developed a very lucrative theory i.e. Analysis of Variance or abbreviated as ANOVA in the year of 1920s to 1930s. Later on Professor Snedecor and many other renounced persons contributed to the development of this technique.

It can be claimed that the analysis of variance is an extremely useful technique concerning researches in the several fields like engineering, education, management, logistics, industry, etc. The first consists in a single factor when examination can be put up and after that multivariate analysis can be imparted.

2. What is Analysis of Variance (ANOVA)?

Analysis of variance is a hypothesis testing technique or statistical tool used to test the egalitarianism of two or more treatment means by investigate the variance of samples that are taken. ANOVA allows one to determine whether the differences between the samples are simply due to sampling error (random error) or whether they are systematic treatment effects that cause the mean in one group to differ from the mean in another.

ANOVA is justified in experimental designs with one dependent variables that is a continuous parametric numerical outcome measure and multiple experimental groups within one or more independent (categorical) variables. In ANOVA terminology, independent variables are reckoned as factors, and groups within each factors are referred to as levels. The array of terms that are part and parcel of ANOVA can be intimidating to the uninitiated, such as source of variation or deviation, sum of squares(SS), mean square (MS), degree of freedom (DOF), f-ratio, multiple comparison procedure(post-hoc tests), effect size, statistical power, etc.

ANOVA can be designed with the help of technique which are as under:

- ✓ One Way ANOVA or single factor ANOVA
- ✓ Two Way ANOVA or double factor ANOVA
- ✓ Latin Square Design ANOVA

3. What is Two Way ANOVA?

Two Way ANOVA is used when the data are classified on the basis of two components. The method can be illustrated with an example like, when a firm of business may have its data for sales off the record on the basis of different salesman persons and also on the basis of sales in different areas(regions). Such a two way ANOVA method may be repeated measurements of each factor or may not have repeated values. The ANOVA technique is little different in case of repeated measurements which we also compute the interaction variation. Now , converse with

- ANOVA technique in context of two way ANOVA when repeated values are there .
- ANOVA technique in context of two way ANOVA when repeated values are not there.

3.1. ANOVA technique in context of Two Way ANOVA design when repeated values are not involved:

Here in this case, repeated values are not inserted. So, we cannot directly compute the Sum of Squares (SS) within samples. Therefore, it is calculated the error variation by subtraction.

The steps involved in the Two Way ANOVA when repeated values are not involved:

I. Firstly, the coded values of individual items in all the samples and find its total which is reckoned its capital as T.

II. Secondly, find out the correction factor which is as follows:

Correction Factor= $\frac{T^2}{n}$,

where n= total number of item in all the samples.

III. Thirdly, find out the square of all the coded values on a case by case basis and then its summation (or total). Subtract the correction factor from this summation to obtain the sum of squares of deviation for total variance which we can write as under:

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

IV. Fourthly, take the total of different columns and then obtain the square of each column total and divide such squared value of each column by the number of items in the

IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 6 Issue 1, January 2019
ISSN (Online) 2348 – 7968
www.ijiset.com

concerning column and take the total of the result thus obtained. Subtract the correction factor from this total to obtain the sum of squares of deviations for variance between columns or SS between columns. Symbolically written as:

SS between columns = $\sum(T_j)^2/n_j$ - $(T)^2/n$

V. Fifthly, take the total of different rows and then obtain the square of each row total and divide such squared values of each row by the number of items in the corresponding row and take the total of the result thus obtained. Subtract the correction factor from this total to obtain the sum of squares of deviation for variance between rows or SS between rows. Symbolically written as:

SS between rows = $\sum(T_i)^2/n_i$ - $(T)^2/n$

VI. Sixthly, Sum of Squares for error variance can be worked out as under:

Total SS – (SS between columns + SS between rows)

VII. Now, degree of freedom can be find out as under:

DOF for variance between columns = (c -1)

DOF for variance between rows = (r -1)

DOF for error variance = (c-1)(r-1)

DOF for total variance = (c×r – 1)

where c is number of columns and r is number of rows

VIII. Finally, the ANOVA table for Two Way ANOVA can be set up which as follows:

Source of variation	Sum of Squares(SS)	Degree of freedom	Mean Square (MS)	F-ratio
SS between columns	$\sum(T_j)^2/n_j - (T)^2/n$	(c-1)	$\frac{SS\ between\ columns}{c-1}$	$\frac{MS\ between\ columns}{MS\ error}$
SS between rows	$\sum(T_i)^2/n_i - (T)^2/n$	(r-1)	$\frac{SS\ between\ rows}{r-1}$	$\frac{MS\ between\ rows}{MS\ error}$
Error deviation	Total SS – (SS between columns and SS between rows)	(c-1)(r-1)	$\frac{SS\ error}{(c-1)(r-1)}$	----------------------------
Total	$\sum X_{ij}^2 - (T)^2/n$	(c×r-1)		----------------------------

3.1.1. Illustration 1, when repeated values are not involved:-

The following table gives the monthly sales (in thousand rupees) of a certain firm in three states by its four salesman. Set up ANOVA table for the information which is given below. Calculate F-coefficients and state whether the difference between sales affected by the four salesman and difference between sales affected in three states are significant. Also taking a significant value of 5%?

States	Salesman			
	A	B	C	D
X	5	4	4	7
Y	7	8	5	4
Z	9	6	6	7

- **Computations for Two Way ANOVA without repeated values**

States	Salesman				Total
	A	B	C	D	
X	5	4	4	7	20
Y	7	8	5	4	24
Z	9	6	6	7	28
Total	21	18	15	18	72

Now, the steps involved are as follows:

i. T =72, n =12

ii. Correction factor = $\frac{(square\ of\ T)}{n}$ = $\frac{72 \times 72}{12}$ = 432

iii. SS total deviation = (25 + 16 +16 + 49 + 49 + 64 + 25 + 16 + 81 + 36 + 36 + 49) – (432)

iv. SS between column deviation = $\frac{(21\times21)}{3} + \frac{(18\times18)}{3} + \frac{(15\times15)}{3} + \frac{(18\times18)}{3} - \frac{(72\times72)}{12}$

= 438-432

= 6

v. SS between rows deviation = $\frac{(20\times20)}{4} + \frac{(24\times24)}{4} + \frac{(28\times28)}{4} - \frac{(72\times72)}{12}$

= (100 + 144 + 196) – 432

= 8

vi. SS error deviation = SS total deviation – (SS between column deviation + SS between row deviation)

= 30 – (6+8)

= 16

Table of ANOVA:

Source of variation	Sum of Squares (SS)	Degree of freedom(DOF)	Mean Square (MS)	F-ratio	5% of F-distribution chart values
SS between columns	6	(4-1) =3	$\frac{6}{3} = 2$	$\frac{(2\times3)}{8} = 3/4$	F(3,6) = 4.76
SS between rows	8	(3-1) = 2	$\frac{8}{2} = 4$	$\frac{(4\times3)}{8} = \frac{3}{2}$	F(2,6) = 5.14
error	16	(4-1)×(3-1) = 6	$\frac{16}{6} = \frac{8}{3}$		
Total	30	4×3 -1 =11	-		

3.2. ANOVA technique in context of Two Way ANOVA design when repeated values are involved:

Here in this case, repeated values are inserted. So, we cannot directly compute the Sum of Squares (SS) within samples. Therefore, it is calculated the error variation by subtraction.

The steps involved in the Two Way ANOVA when repeated values are not involved:

I. Firstly, the coded values of individual items in all the samples and find its total which is reckoned its capital as T.

II. Secondly, find out the correction factor which is as follows:

Correction Factor= $\frac{Square\ of\ T}{n}$, where n= total number of item in all the samples.

III. Thirdly, find out the square of all the coded values on a case by case basis and then its summation (or total). Subtract the correction factor from this summation to obtain the sum of squares of deviation for total variance which we can write as under:

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

IV. Fourthly, take the total of different columns and then obtain the square of each column total and divide such squared value of each column by the number of items in the concerning column and take the total of the result thus obtained. Subtract the correction factor from this total to obtain the sum of squares of deviations for variance between columns or SS between columns. Symbolically written as:

SS between columns = $\sum(T_j)^2/n_j - (T)^2/n$

V. Fifthly, take the total of different rows and then obtain the square of each row total and divide such squared values of each row by the number of items in the corresponding row and take the total of the result thus obtained. Subtract the correction factor from this total to obtain the sum of squares of deviation for variance between rows or SS between rows. Symbolically written as: SS between rows = $\sum(T_i)^2/n_i - (T)^2/n$

VI. Sixthly, determine the sum of squares within the samples(error variation)

VII. Seventhly, Sum of Squares for interrelationship(or interaction) variation can be worked out as under:

Total SS – (SS between columns + SS between rows + SS within samples)

VIII. Now, degree of freedom can be find out as under:

DOF for variance between columns = (c -1)

DOF for variance between rows = (r -1)

DOF for error variance = Total number of sample(n) – Half of the total number of sample(n)

DOF for interaction variation = [DOF of total variance – (DOF of columns + DOF of rows + DOF of error]

DOF for total variance = [number of columns(c) × total number of rows(r) – 1]

ISSN (Online) 2348 – 7968
www.ijiset.com

where c is the number of columns and r is the number of rows

IX. Finally, the ANOVA table for Two Way ANOVA can be set up which as follows:

Source of variation	Sum of Squares(SS)	Degree of freedom	Mean Square (MS)	F-ratio
SS between columns	$\sum(T_i)^2/n_i - (T)^2/n$	(c-1)	$\frac{SS\ between\ columns}{c-1}$	$\frac{MS\ between\ columns}{MS\ error}$
SS between rows	$\sum(T_i)^2/n_i - (T)^2/n$	(r-1)	$\frac{SS\ between\ rows}{r-1}$	$\frac{MS\ between\ rows}{MS\ error}$
Interrelationship	Total SS – (SS between columns + SS between rows + SS within samples)	[DOF of total variance – (DOF of columns + DOF of rows + DOF of error]	$\frac{SS\ of\ Interrelationship}{Degree\ of\ Freedom\ of\ Interrela}$	$\frac{MS\ between\ interrelations}{error\ deviation}$
Within error deviation	Within sales error can be calculated	Total number of sample(n) – Half of the total number of sample(n)	$\frac{SS\ error}{DOF}$	----------------------------
Total	$\sum X_{ij}^2 - (T)^2/n$	(c×r-1)		-----------------------------

3.2.1. Illustration 2 when repeated values are involved:-

The following table gives the monthly sales (in thousand rupees) of a certain firm in three states by its four salesman. Set up ANOVA table for the information which is given below. Calculate F-coefficients and state whether the difference between sales affected by the four salesman and difference between sales affected in three states are significant. Also taking a significant value of 5%?

States	**Salesman**			
	A	B	C	D
X	5 3	4 5	4 9	7 8
Y	7 3	8 8	5 7	4 5
Z	9 5	6 4	6 3	7 1

➢ **Computations for Two Way ANOVA with repeated values**

States	**Salesman**				**Total**
	A	B	C	D	
X	53	45	49	78	45
Y	73	88	57	45	47
Z	95	64	63	71	41
Total	32	35	34	32	133

Now, the steps involved are as follows:

i. T = 133, n = 24

ii. Correction Factor = $\frac{(square\ of\ T)}{n}$ = $\frac{(133\times133)}{24}$

= 737.04

= 737

iii. SS total deviation = [25 + 9 + 16 + 25 + 16 + 81 + 49 + 64 + 49 + 9 + 64 + 64 + 25 + 49 + 16 + 25 + 81 + 25 + 36 + 16 + 36 + 9 + 49 + 1] – $\frac{(133\times133)}{24}$

= 823 - 737

= 86

iv. SS between columns (i.e. between salesman) deviation

= [$\frac{(32\times32)}{6}$ + $\frac{(35\times35)}{6}$ + $\frac{(34\times34)}{6}$ + $\frac{(32\times32)}{6}$] - $\frac{(133\times133)}{24}$

= (170.6 + 204.1 + 192.6 + 170.6) – 737.04

= 737.9 – 737.04

= 0.86

v. SS between rows (i.e. between states) deviation

= [$\frac{(45\times45)}{8}$ + $\frac{(47\times47)}{8}$ + $\frac{(41\times41)}{8}$] - $\frac{(133\times133)}{24}$

= (253.1 + 276.1 + 210.1) – 737.04

= 739.3 – 737.04

= 2.26

vi. SS within sales (i.e. error) deviation

= (1 + 1 + 0.25 + 0.25 + 6.25 + 6.25 + 0.25 + 0.25 + 4 + 4 + 0 + 0 + 1 + 1 + 0.25 + 0.25 + 4 + 4 + 1 + 1 + 2.25 + 2.25 + 9 + 9)

= 58.5

vii. SS for interrelationship variation

= 86 – (0.86 + 2.26 + 58.5)

= 24.38

Table of ANOVA:

Source of variation	Sum of Squares (SS)	Degree of freedom (DOF)	Mean Square (MS)	F- ratio	5% of F- distribution chart values
SS between columns	0.86	(4 -1) = 3	$\frac{0.86}{3}$ = 0.28	$\frac{0.28}{4.87}$ = 0.057	F(3,12) = 3.49
SS between rows	2.26	(3-1) = 2	$\frac{2.26}{2}$ = 1.13	$\frac{1.13}{4.87}$ = 0.23	F(2,12) = 3.88
SS of Interrelationship	24.38	4	$\frac{24.38}{4}$ = 6.0	$\frac{6.0}{4.87}$ = 1.23	F(4,12) = 3.26
SS within sales(error)	58.5	(24-12) = 12	$\frac{58.5}{12}$ = 4.87		
Total	86	(4×6 -1) = 23			

4. Conclusion:

ANOVA technique have been used for ages into various disciplines. But, now a days, it is used rapidly because of multivariate analysis is observed easily without any obstacles which allows more complete examination of the object. As it was shown in this article, the two way ANOVA technique is appropriate to consider the problems associated the domain of marketing (or sales). Therefore, the applications in this paper with the high potential analysis of variance carries.

Reference:

[1] Niedoba T., and Pieta P., 2016, “Application of ANOVA in Mineral Processing”, Mining Science, Vol. 23, pp. 43-54.12

[2] Steven F. S., “Analysis of Variance: The Fundamental Concepts”, The Journal of Manual and Manipulative Therapy, Vol. 17, pp. 27-38.

[3] Mouhamadou T S., 2014, “Using ANOVA to examine the Relationship between Safety & Security and Human Development”, Journal of International Business and Economics, Vol. 2 (Issue.4), pp. 101-106.

[4] Ostertagova E., and Ostertag O., 2013, “Methodology and Application of One Way ANOVA”, American Journal of Mechanical Engineering, Vol. 1 (Issue.7), pp. 256-261.

Books: Andrew Rutherford, (2001), “Introducing ANOVA and ANOCOVA- a

GLM Approach", First Edition, SAGE Publications.

Book Chapter: C.R.Kothari, (2004), Analysis of Variance and Co-variance, Research Methodology: Methods and Applications, New Age International Publishers, pp. 256-282.