iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 3 **Issue:** IX **Month of publication:** September 2015

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Some Pointers on One Way ANOVA in SPSS

Sahana Prasad
Associate Professor, Department of Statistics, Christ University, Bangalore, India

Abstract— ANOVA is an important tool used in varied fields. This acronym refers to "analysis of variance" and is a widely used statistical procedure. It is applicable in those situations where there are two or more groups of data and we need to test the degree to which these groups differ in an experiment. A higher variance generally implies that there was a significant finding from the research/ experimentation. There are many types of ANOVA, each applied to different situations. There are a few important assumptions under which ANOVA can be used and it is essential to test them all on the data. This paper gives a brief introduction to the procedure as well discusses some applications using SPSS.
Keywords— ANOVA, one way ANOVA, applications, assumptions, SPSS

I. INTRODUCTION

ANOVA Means "Analysis Of Variance". It is a statistical process for analysing the amount of variance that is contributed to a sample by different factors. It was initially derived by R. A. Fisher in 1925. In statistical experiments, we first formulate a null hypothesis and an alternative hypothesis. A null hypothesis is the assumption that there will be no differences between groups that are tested and therefore, no significant results will be revealed. Alternative hypothesis is the hypothesis stating that there will be a difference between groups. ANOVA will help to test this assumption. It can be regarded as an extension of paired sample t-test and there are a few assumptions to be tested in the data before performing ANOVA. The main aim of one-way -ANOVA is to determine whether there exist any significant differences between the means of two or more independent groups. It is also is important to realize that the one-way ANOVA will not be able to explain which specific groups were significantly different from each other. It is a procedure in which the difference between several sample means is tested simultaneously. Variance is the arithmetic average of the square of deviations of data from their means. Variance has additive property, which is not possessed by standard deviation. Thus, variance can be added up and also, broken down into various components. ANOVA is thus a tool for breaking up the total variance of a large sample or population which consists of many groups. The variance is divided into two types:

"within-groups" variance which is the average of the members of each group around their respective group means

"between-groups" variance which is the variance of the group means around the total or grand mean of all groups.

ANOVA is favoured by social scientists because it can be applied to situations where experimental variables have been held at constant levels but others may not have been. In social sciences, the researcher may not have freedom to manipulate his subjects or their environment. In ANOVA, we identify which of the several variables are truly independent.

II. WHY IS ANOVA NEEDED

Sir R. A. Fisher (1935) clearly explained the relationship among the mean, the variance, and the normal distribution: "The normal distribution has only two characteristics, its mean and its variance. The mean determines the bias of our comparison. This involves lot of time as well as computation. As the sample sizes increase, the number of computations will also go up. Secondly, even after all the lengthy computations; there may not be a significant difference between various pairs, which will not help the researcher. Thirdly, it may turn out that some maybe significant at 0.01 level, some at 0.05 level while some may not be significant at any level. This again brings the researcher to a dilemma. Thus, a single composite test is required. Fourthly, the differences between means are not independent of each other. This means, if we know the difference between stream A and B, as well as A and C, we can find the difference between B and C. This shows that one difference is not independent of the other two differences. But both z and t tests are used to find differences between pairs of means of two independent random samples. Thus, z and t-tests are not feasible tools here.

A. Assumptions In ANOVA

No statistical test can be arbitrarily used and the researcher has to ensure that a specific set of conditions must be met. Such conditions are known as "Model Assumptions". For ANOVA, these are:

Independence – which means selection of a participant should not be dependent on selection of any other participant.

The most important issue of independence is that "observations within or between groups are not paired, dependent, correlated, or associated in any way" (Glass & Hopkins, 1996, p. 295).

Estimates of Type I and type II errors are not accurate when this condition of independence is violated.

Randomness in selection and assignment is important too, as this provides an equal chance for all units to be selected. It may not always be possible in social sciences as certain characteristics like gender cannot be assigned randomly. All units must be measured only once. For e.g. If a student has chosen social work and statistics for his study, he should not be measured twice. It would be better to omit such units from the study.

Normality- If we consider a normal sample distribution, the means and variances over repeated samples will be uncorrelated, and therefore, statistically independent. The formulae used to compare groups are based on the mean, median and mode being approximately equal, which is a property of normal distribution.

Dependent variable should be measured at the interval or ratio level

Independent variable should consist of two or more categorical, independent groups.

There should be no significant outliers.

Caution: The sample chosen plays a very important role. If the sample has people from higher caste, the results on, say reservations for lower castes may not be valid. The conclusions must be restricted to the chosen sample alone. Most research designs address the problems of internal validity rather than generalization.

III. ANOVA IN SPSS

Formulate null an alternative hypothesis.
Check data requirements by studying the data and checking if (iv) and (v) listed above are satisfied.
Check for outliers using "EXPLORE" option.
In "EXPLORE", check normality plots option to check for normality.
Using the output of Shapiro Wilk test, confirm normality.
Homogeneity of variances can be checked using Levenes's test.
Use the option of one way ANOVA option and feed in the required details.
Draw conclusions from the analysis.
Post hoc tests are run in order to ascertain where the differences have occurred between groups. They are used only when there is a significant difference in group means .They are called "Post-hoc" as they are performed after the event.
(See Section V)

IV. SOME CASE STUDIES

A. Case Study 1

A psychologist predicts that students will learn most effectively with a constant background sound, as opposed to an unpredictable sound or no sound at all. She randomly divides twenty-four students into three groups of eight. All students study a passage of text for 30 minutes. Those in group 1 study with background sound at a constant volume in the background. Those in group 2 study with noise that changes volume periodically. Those in group 3 study with no sound at all. After studying, all students take a 10 point multiple choice tests over the material. Their scores are as follows:

Group	Test scores							
Constant sound	7	4	6	8	6	6	2	9
Random sound	5	5	3	4	4	7	2	2
No sound	2	4	7	1	2	1	5	5

Suitable tool: One way ANOVA

Reason: The "sound" is the only variable here which may be affecting the scores. For reasons given in Case study 1, ANOVA is the

suitable choice.

Solution: Ho: There is no significant difference between test scores of various groups.

H1: There is a significant difference between test scores of various groups.

The ANOVA table is given below:

Sources of Variation	Sum of Squares	Degrees of freedom	Mean sum of squares	F
Within groups	87.88	21	4.18	3.59
Between groups	30.08	2	15.04	

Table value of F at (2, 21) at 0.05 = 3.4668

Conclusion: The psychologist can conclude that hypothesis H1 may be supported. The means are as he/she predicted and the constant music group has the highest score.

Caution: However, the significant F only indicates that at least two means are significantly different from one another, but the researcher will not know which specific mean pairs significantly differ until a post-hoc analysis like Tukey's HSD is conducted.

B. Case Study 2

A study was reported by Hogg and Ledolter (1987, of bacteria counts in shipments of milk. There were five shipments and for each shipment bacteria counts were made from 6 randomly selected cartons of milk. The question to be answered is whether some shipments have higher bacteria counts than others. The data are shown below,

Bacteria count data, shipments 1-5, and samples 1-6

	1	2	3	4	5
1	24	14	11	7	19
2	15	7	9	7	24
3	21	12	7	4	19
4	27	17	13	7	15
5	33	14	12	12	10
6	23	16	18	18	20

Suitable tool: One way ANOVA

Reason: The "bacterial count" is the only variable here which may be affecting the cartons of milk. For reasons given in Case study 1, ANOVA is the suitable choice.

Solution: Ho: There is no significant difference between milk cartons with respect to bacterial count.

H1: There is a significant difference between milk cartons with respect to bacterial count.

The ANOVA table is given below:

Sources of Variation	Sum of Squares	Degrees of freedom	Mean sum of squares	F
Within groups	803	4	200.75	9.01
Between groups	557.17	25	22.287	

Table value of F at (4, 25) at 0.05 = 2.7587

Volume 3 Issue IX, September 2015
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

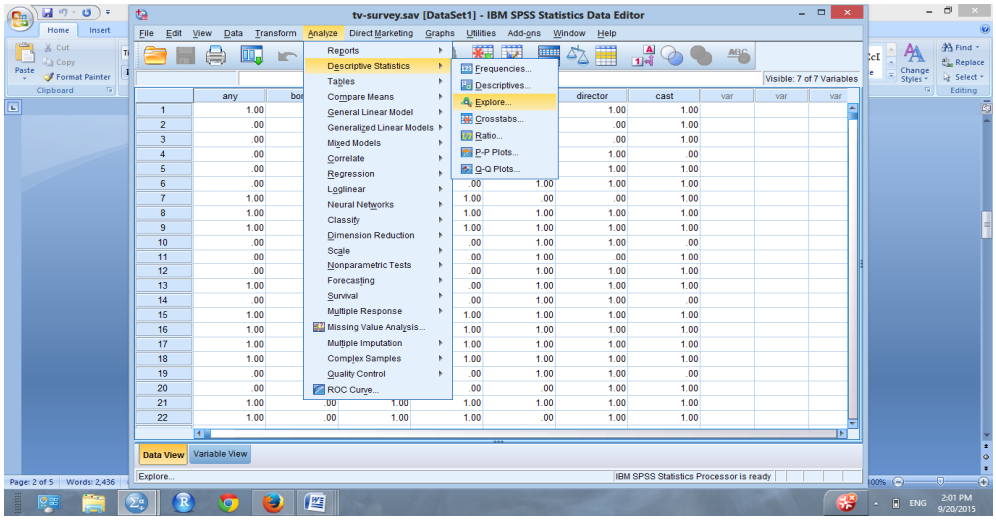
Conclusion: The researcher can conclude that hypothesis H1 may be supported. There is a significant difference in the bacterial counts which is definitely not due to chance.

IV. RESULTS AND DISCUSSIONS

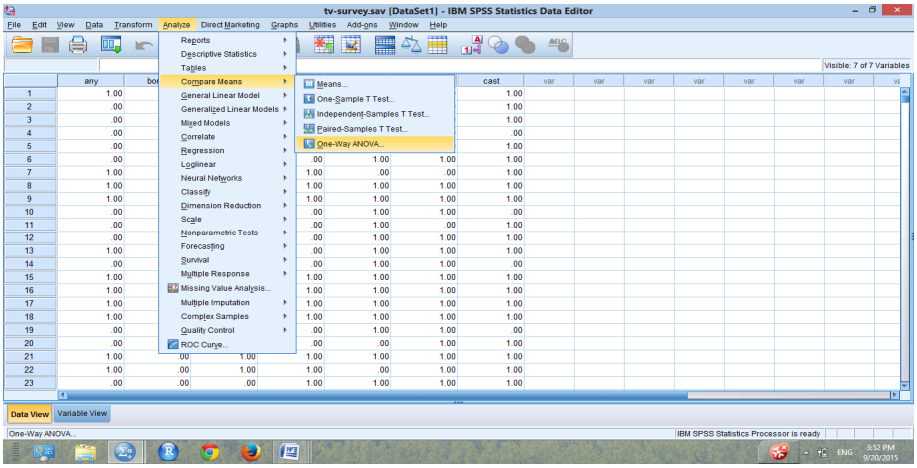
In analysis of variance, when the observed value of F exceeds the expected theoretical value, it means the samples are from different populations. It also tells us that there are differences between treatments in the experiment as a whole but it does not tell us which treatments differ from one another. to check which pair has different means, we apply student's t-test assuming that all populations share a common variance. the critical difference method is then used to find the pairs.

V. SPSS COMMANDS

ANOVA for "EXPLORE" option

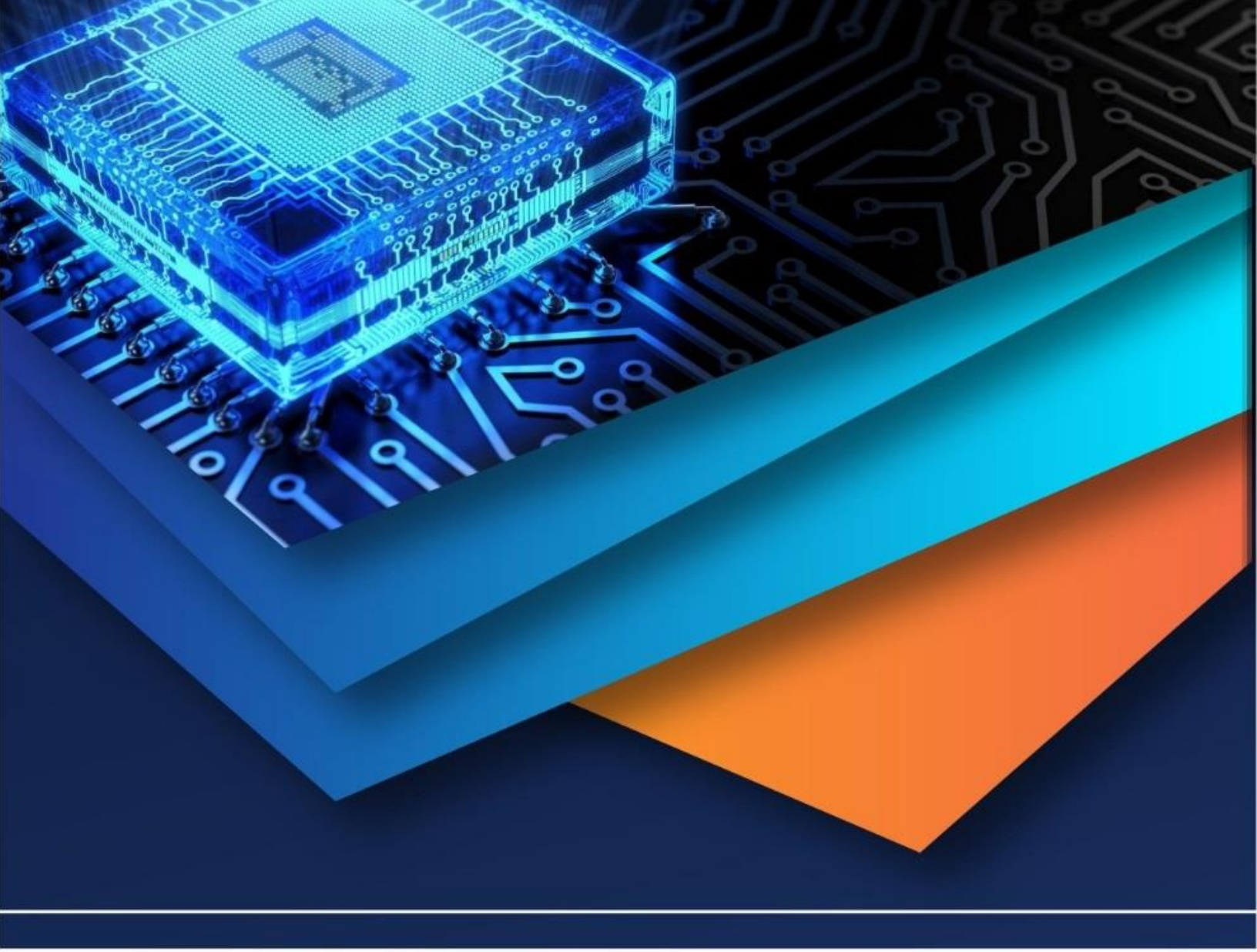


ANOVA for "One way"



REFERENCES

[1] Sahana Prasad, monograph on Application of Statistical tools in Social Sciences research, published by Christ University, Bangalore, India
[2] https://statistics.laerd.com/features-spss-procedure.php
[3] Healey, J. (n.d.). Statistics- A tool for Social Research. Newport University.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY