Analysis Of Variance (ANOVA)

Analysis of Variance (ANOVA) is a parametric statistical technique used to compare datasets. This technique was invented by R.A. Fisher, and is thus often referred to as Fisher's ANOVA, as well. It is similar in application to techniques such as t-test and z-test, in that it is used to compare means and the relative variance between them. However, analysis of variance (ANOVA) is best applied where more than 2 populations or samples are meant to be compared. An ANOVA test is a way to find out if survey or experiment results are significant. In other words, they help you to figure out if you need to reject the null hypothesis or accept the alternate hypothesis. Basically, you're testing groups to see if there's a difference between them. **Analysis of variance (ANOVA)** is a collection of statistical models and their associated estimation procedures (such as the "variation" among and between groups) used to analyze the differences among group means in a sample.

The use of this parametric statistical technique involves certain key assumptions, including the following:

1. **Independence of case:** Independence of case assumption means that the case of the dependent variable should be independent or the sample should be selected randomly. There should not be any pattern in the selection of the sample.

2. **Normality:** Distribution of each group should be normal. The Kolmogorov-Smirnov or the Shapiro-Wilk test may be used to confirm normality of the group.

3. **Homogeneity:** Homogeneity means variance between the groups should be the same. Levene's test is used to test the homogeneity between groups.

If particular data follows the above assumptions, then the analysis of variance (ANOVA) is the best technique to compare the means of two, or more, populations.

Analysis of variance (ANOVA) has three types:

One way analysis: When we are comparing more than three groups based on one factor variable, then it said to be one way analysis of variance (ANOVA). For example, if we want to compare whether or not the mean output of three workers is the same based on the working hours of the three workers.

Two way analysis: When factor variables are more than two, then it is said to be two way analysis of variance (ANOVA). For example, based on working condition and working hours, we can compare whether or not the mean output of three workers is the same.

K-way analysis: When factor variables are k, then it is said to be the k-way analysis of variance (ANOVA).

Key terms and concepts:

Sum of square between groups: For the sum of the square between groups, we calculate the individual means of the group, then we take the deviation from the individual mean for each group. And finally, we will take the sum of all groups after the square of the individual group.

Sum of squares within group: In order to get the sum of squares within a group, we calculate the grand mean for all groups and then take the deviation from the individual group. The sum of all groups will be done after the square of the deviation.

F –ratio: To calculate the F-ratio, the sum of the squares between groups will be divided by the sum of the square within a group.

Degree of freedom: To calculate the degree of freedom between the sums of the squares group, we will subtract one from the number of groups. The sum of the square within the group's degree of freedom will be calculated by subtracting the number of groups from the total observation.

Significance: At a predetermine level of significance (usually at 5%), we will compare and calculate the value with the critical table value. Today, however, computers can automatically calculate the probability value for F-ratio. If p-value is lesser than the predetermined significance level, then group means will be different. Or, if the p-value is greater than the predetermined significance level, we can say that there is no difference between the groups' mean.

Analysis of variance (ANOVA) in SPSS: In SPSS, analysis of variance (ANOVA) can be performed in many ways. We can perform this test in SPSS by clicking on the option "one way ANOVA," available in the "compare means" option. When we are performing two ways or more than two ways analysis of variance (ANOVA), then we can use the "univariate" option available in the GLM menu. SPSS will give additional results as well, like the partial eta square, Power, regression model, post hoc, homogeneity test, etc. The post hoc test is performed when there is significant difference between groups and we want to know exactly which group has means that are significantly different from other groups.
Extension:

MANOVA: Analysis of variance (ANOVA) is performed when we have one dependent metric variable and one nominal independent variable. However, when we have more than one dependent variable and one or more independent variable, then we will use multivariate analysis of variance (MANOVA).