Things to Remember - ANOVA

In order to compare multiple groups at once, we can look at the ANOVA, or Analysis of Variance. Unlike the t-test, it compares the variance within each sample relative to the variance between the samples.The test for differences between more than two sample means is accomplished with F distribution.

Why ANOVA ?

Very often an analyst wants to compare two population means in certain situations. For example, a human resource manager may be interested in the aptitude test scores from a standard test for executives having four types of background. In order to make such comparison usually t distribution is used by looking at samples two at a time and then comparing the means. Though this method is feasible it turns out to be inefficient as while using this method we will find a significant difference exists between at least one pair of means when there are enough identical populations.

With this problem we cannot trust the level of significance and require a technique that will enable the analyst to test the hypothesis of equality of more than two population means that will overcome the deficiencies of t test . It is in this ANOVA - Analysis of Variance plays an important role.

For ANOVA, the null hypothesis is taken that all means are equal while the alternative hypothesis is not all of them are equal or at least two of them are different

- H0 :μ1 = μ2 = μ3 = ... μk - All population means are equal
- H1 : Not all of the population means are equal -(For at least one pair, the population means are unequal)
- One way ANOVA examines the effect of one variable only.
- It consists of obtaining the ratio of two variances - Variance between the samples and variance within the sample. This ratio is expressed as F = Variances between samples / variance within the sample.
- The calculated value of F is compared with the critical value obtained from the F values.
- The two variances are calculated by the Total Sum of Squares ≡ Treatment Sum of Squares + Error Sum of Squares.
- The word treatment is generic and as such may denote different methods, machines, different advertisement copy platforms, different strategies, different brands and the like
- The variation in sum of squares of the response variable (dependent variable) is caused only by treatment and any thing unexplained by the treatment is attributed to error term.

Assumptions of one way ANOVA

- Normality – That each sample is taken from a normally distributed population
- Sample independence – that each sample has been drawn independently of the other samples
- Variance Equality – That the variance of data in the different groups should be the same
- Your dependent variable – here, “weight”, should be continuous – that is, measured on a scale which can be subdivided using increments (i.e. grams, milligrams)

This file is meant for personal use by sudhirjoon05@gmail.com only.
Sharing or publishing the contents in part or full is liable for legal action.

Two way ANOVA

Its is mere extension of one way ANOVA.in the two-way ANOVA each sample is defined in two ways, and resultantly put into two categorical groups for eg.

- Human Resources Department of a company desires to know if occupational stress varies according to age and gender.
    - The variable of interest is therefore **occupational stress** as measured by a scale.
    - There are two **factors** being studied - age and gender.

Assumptions of two way ANOVA

- Your dependent variable – that is, measured on a scale (continuous) which can be subdivided using increments (i.e. grams, milligrams)
- Your two independent variables – should be in categorical, independent groups.
- Sample independence – that each sample has been drawn independently of the other samples
- Variance Equality – That the variance of data in the different groups should be the same
- Normality – That each sample is taken from a normally distributed population

Summary of one way and two way Anova

	One way ANOVA	Two way ANOVA
Definition	Allows to make comparison between the means of three or more groups of data	Allows to make comparison between means of three or more groups of data, where two independent variables are considered
Number of independent variables	One	Two
What is being compared	The means of three or more groups of an independent variable on a dependent variable.	The effect of multiple groups of two independent variables on a dependent variable and on each other.
Number of groups of samples	Three or more.	Each variable should have multiple samples.

Summary of tests

This file is meant for personal use by sudhirjoon05@gmail.com only.
Sharing or publishing the contents in part or full is liable for legal action.

	Types of Data	
Type of Analysis	Numerical	Categorical
Compare two population means	t tests for the difference in the means of two independent populations	Z test for the difference between two proportions
Compare more than two populations	One way ANOVA	

This file is meant for personal use by sudhirjoon05@gmail.com only.
Sharing or publishing the contents in part or full is liable for legal action.