

# ANOVA

## Overview

The purpose of one-way ANOVA is used to test for differences on means among two or more independent groups. In practice the One-way ANOVA is used to test for differences among three or more groups, with the two-group case relegated to the t-test which is a special case of the ANOVA. The relation between ANOVA and t is given as  $F = t^2$ .

## Hypothesis

Null  $H_0: \mu_i = \mu_j$  for all  $i \neq j$   
Alternative  $H_1: \mu_i \neq \mu_j$  for some  $i \neq j$

## Assumptions

*Random samples*-The samples are simple random samples or random assignments

*Independence* - The samples are independent of each other

*Normality* - the populations have normal distributions

*Homogeneity of variances*- the variance of data in groups should be the same

## Logic of ANOVA

At the heart of ANOVA is the fact that variances can be divided up, that is, partitioned. The basic idea is to compare the variance due to difference in means, called mean square between groups (MSBG) vs. the variance due to true random error called mean square inside groups (MSWG). If they are similar, there is no evidence to suggest a difference between the group means (fail to reject null), if they are different, there is evidence in favor of a difference between means (reject null). This comparison of variances is accomplished by using the F test.

## Procedure

We partition the total sum of squares into components related to the effects used in the model, that is; the variation due to difference between the groups, called sum of squares between groups (SSBG) and the natural random variation inside a group called sum of squares within groups (SSWG). The number of degrees of freedom (abbreviated df) can be partitioned in a similar way for each component. The mean squares (MS) are computed dividing the sum of squares (SS) by the corresponding degrees of freedom. The F statistics is the ratio of MSBG to MSWG.

**Example** (Page 652, problem 14, Triola, 2007)

Consider  $p = \#$  of groups

$n =$  number of observations per group

$nT =$  total number of observations

## ANOVA TABLE

Source	DF	SS	MS	F	P
Factor (BG)	$p-1=2$	138.7	69.4	4.05	0.031
Error (WG) equal n's	$p(n-1)=24$	411.1	17.1		
Total	$nT-1$ or $np-1=26$	549.9			

## Where:

$MS = SS/df$ ,  $F = MSBG/MSWG$

$p =$  pvalue or level of significance