

One way Analysis of Variance (1-Way ANOVA) can be used to compare 2 or more population means when samples are independent:

Given  $K$  independent samples ( $K \geq 2$ ), test the null hypothesis

$H_0: \mu_1 = \mu_2 = \dots = \mu_K$  vs, the alternative hypothesis  $H_1$ : not all population means are equal

$\{(X_{11}, \dots, X_{1,n_1}), \dots, (X_{K1}, \dots, X_{K,n_K})\}$

**ANOVA** analyses total variance in the combined data to come up with a test (the F-test) for comparison of  $K$  means.


Linear Regression (Simple or Multiple) fits a linear model to a response variable  $Y$  as a function of the predictor(s) in the data. ANOVA can be thought of as regression when predictor(s) are categorical.


Linear model for both ANOVA and regression has the form

$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_K X_K + e$ ,  $e$  is random error assumed to be normally distributed with 0 mean and a common variance  $\sigma^2$ .

One should always verify that the errors (residuals) are normally distributed; one common statistical test for normality is the Shapiro Test; a graphical verification of normality is the Q-Q plot.


Watch the video lecture SLR and MLR V.mp4.

⋮  6\_3 Testing Hypotheses Involving Population Proportion V.pptx

⋮  one and two sample t-tests and CIs.mp4

⋮  testing normality of 2 samples.txt

⋮  SLR and MLR V.mp4 

⋮  8a1 Correlation and Simple Linear Regression V.pptx

For the data file ANOVA Golf Example 1.csv, determine if Brand has a significant effect on Distance (1-Way ANOVA and Regression methods).

For the data file 2 way ANOVA Golf Example.csv, determine if Golfer and Brand have

significant effects on Distance (2-Way ANOVA and Regression methods). Test normality of residuals for each of the above two problems and methods.