

Hypothesis Test

- ~ Null hypothesis (H_0) : ~ There is no difference
- ~ Alternate/Research Hypothesis (H_1) : ~ There is some difference

Type of test :

~ P value	~ Population < 30	~ S.D is not known	one sample with two category
~ T-test	↓	~ S.D is known	one sample with two category & large population
~ Z-test	Population ≥ 30		one sample with more than two categories
~ Anova test	Extension of T-test		for two sample test
~ Chi square			for categorical variables

parametric (Population Parameter) → (Mean, Variance, mode, S.D)

↓ T-test ↓ Z-test ↓ F-test ↓ Anova test

Non-parametric (Sample Parameter) → (Mean, mode, Variance, S.D)

↓ Chi square test

$$Z\text{-test} = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$
 → \bar{X} :- Sample Mean
 → μ_0 :- Population Mean
 → σ :- Standard deviation (S.D)
 → n :- No of observation

$$T\text{-test} = \frac{\text{Difference b/w mean}}{\text{Variance / Sample size}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$\text{Anova-test} = \frac{MST}{MSE}$$

category	Continuous
chi-square	T-test Anova

$$\text{Chi-square } (\chi^2) = \frac{\sum \frac{O^2}{E}}{\sum \frac{O^2}{E}} (n-1)$$

Continuous	Logistic Regression	correlation -1 to +1
------------	---------------------	----------------------

~ P value is Significance or Probability of Null Hyp

Table Value / P value / Significance

for 5% → 1.96
 for 1% → 2.58

P value < computed value
 Reject the Null Hypothesis

P value > computed value
 Accept the Null Hypothesis