Statistical Tests

Contionous Variable

- 1 sample Test
  - 1 sample Z test
  - 1 sample T test
- 2 sample Test
  - 2 sample Test for equal variance
  - 2 sample Test for unequal variance
- Anova test

Discrete Variable

- 1 proportion test
- 2 proportion test
- Chisquare test

Steps involved in Statistical Tests (Hypothesis testing)

Step 1 : Formulate the Null Hypothesis (H0) and Alternate Hypothesis (H1)

Step 2 : Level of Significance (alpha)

Step 3 : Check for Normality of given data (for Continous data only) --> 1 sample
Check for variances are equal (2 or more samples)

Step 4 : Select the Statistical Test and calculate p value

Step 5 : Based on p value , conclude the Hypothesis test

if p<= alpha (p low Null go) ---> Reject H0

if p>alpha (p high Null fly) ---> do not Reject H0

Check for Normality

Option 1 - Skewness:

- if the skewness is between -1 to +1 --> then, it said to be a normal distribution
- if the skewness is less than -1 --> then, it said to be a left skewed distribution
- if the skewness is greater than +1 --> then, it said to be a right skewed distribution

Option 2 - Density Curve:

- if the density curve is symmetrical --> then, it said to be a normal distribution
- if the density curve is non symmetrical --> then, it said to be a Skewed distribution

Option 3 - Shapiro test

- H0 : Data is normal --> if p>alpha, then Data is normal
- H1 : Data is not normal --> if p<=alpha, then Data is not normal

Fabric Data

The length of 25 samples of a fabric are taken at random. Mean and standard deviation from the historic 2 years study are 150 and 4 respectively. Test if the current mean is greater than the historic mean. Assume α to be 0.05

Sol: Given 1 Sample, data is Normal & Population std. dev. is known → **1-Sample Z test**

Bolt Diameter

The mean diameter of the bolt manufactured should be 10mm to be able to fit into the nut. 20 samples are taken at random from production line by a quality inspector. Conduct a test to check with 95% confidence that the mean is not different from the specification value.

Sol: Given 1 Sample, Data is Normal & Population std. dev. is unknown → **1-Sample T test**

Marketing Strategy

A financial analyst at a Financial institute wants to evaluate a recent credit card promotion. After this promotion, 500 cardholders were randomly selected. Half received an ad promoting a full waiver of interest rate on purchases made over the next three months, and half received a standard Christmas advertisement. Did the ad promoting full interest rate waiver, increase purchases?

Sol: Given 2 Samples, Both are Normal and 2 sample variances are equal → **2-Sample T test**

Contract Renewal

A marketing organization outsources their back-office operations to three different suppliers. The contracts are up for renewal and the CMO wants to determine whether they should renew contracts with all suppliers or any specific supplier. CMO want to renew the contract of supplier with the least transaction time. CMO will renew all contracts if the performance of all suppliers is similar.

Sol: Given 3 Samples, All are Normal, 3 Variances are equal → **One way ANOVA Test**

Football Coach

The people carry out a poll to find the acceptability of new football coach. It was decided that if the support rate for the coach for the entire population was truly less than 25%, the coach would be fired. 2000 people participated and 482 people supported the new coach. Conduct a test to check if the new coach should be fired with 95% level of confidence.

Sol: Given 1 Sample proportion → **1 Proportion Test**

Johnnie Talkers

Johnnie Talkers soft drinks division sales manager has been planning to launch a new sales incentive program for their sales executives. The sales executives felt that adults (>40 yrs) won't buy, children will & hence requested sales manager not to launch the program. Analyze the data & determine whether there is evidence at 5% significance level to support the hypothesis

Sol: Given 2 proportions → **2 Proportion Test**

Bahaman Research

Baha ManTech Research Company uses 4 regional centers in South Asia (India, China, Srilanka and Bangladesh) to input data of questionnaire responses. They audit a certain % of the questionnaire responses versus data entry. Any error in data entry renders it defective. The chief data scientist wants to check whether the defective % varies by country. Analyze the data at 5% significance level and help the manager draw appropriate inferences. ['1' means not defectives & '0' means defective]

Sol: Given >2 proportions → **Chi-Square Test**