Selection, Validation, and Reporting Statistical Tests

There are many more, we talk about the most common ones here.

1

Define the Hypothesis

The research question or hypothesis should be clearly defined and specific to enable the selection of the appropriate statistical test.

2

Determine the Type of Data

Identify the type of data being used in the study, whether it is categorical or numerical.

3

Determine Number of Groups

This step involves determining the number of groups that will be compared in the study Is it two or more than two?

The groups may be independent, such as two different treatment groups, or dependent, such as pre-and post-treatment groups.

4

Determine the Nature of the Dependent Variable

The dependent variable is the variable that is being measured in the study. It may be continuous, dichotomous, or discrete count.

5

Data Assumption Tests

NORMALITY TEST - Shapiro-Wilk test, Kolmogorov-Smirnov test
VARIANCE TEST - Levene's test, F-test
CORRELATION STRENGTH TEST - Pearson's correlation coefficient
TIME-BASED TEST - Stationary test
SURVIVABILITY TEST - Log-rank test
SET CORRELATION TEST - Bartlett's test of sphericity

6a

Choose the Statistical Test: t-test

Purpose:

To compare the means of two independent groups with a continuous dependent variable.

Assumptions:

Normality, equal variance, and independence of observations.

Tests Required:

Shapiro-Wilk test, Levene's test.

6b

Choose the Statistical Test: ANOVA

Purpose:

To compare the means of more than two independent groups with a continuous dependent variable.

Assumptions:

Normality, equal variance, and independence of observations.

Tests Required:

Shapiro-Wilk test, Levene's test.

6c

Choose the Statistical Test: Chi-squared

Purpose:

To compare the proportions of two or more groups with a dichotomous or categorical dependent variable.

Assumptions:

Categorical or dichotomous data and expected frequency in each cell >= 5.

Tests Required:

None

6d

Choose the Statistical Test: Correlation Analysis

Purpose:

To determine the strength and direction of the linear relationship between two continuous variables.

Assumptions:

Linearity, normality, and continuous data.

Tests Required:

Shapiro-Wilk test, Pearson's correlation coefficient.

6e

Choose the Statistical Test: Regression Analysis

Purpose:

To determine the relationship between one dependent variable and one or more independent variables.

Assumptions:

Linearity, normality, and homoscedasticity of residuals.

Tests Required:

Shapiro-Wilk test, scatter plot, residual plot, F-test.

6f

Choose the Statistical Test: Time Series Analysis

Purpose:

To model and forecast time-dependent data.

Assumptions:

Time dependency and adherence to a specific pattern.

Tests Required:

Autocorrelation plot, stationarity tests.

6g

Choose the Statistical Test: Survival Analysis

Purpose:
To analyze the time to an event, such as death or failure.

Assumptions:
Time dependency, adherence to a specific hazard function, and independence of censoring.

Tests Required:
Kaplan-Meier plot, Log-rank test.

6h

Choose the Statistical Test: Cluster Analysis

Purpose:
To group similar data points together into clusters.

Assumptions:
Appropriate distance metric based on the data type.

Tests Required:
None

6i

Choose the Statistical Test: Factor Analysis

Purpose:
Identify underlying factors and reduce the number of variables by identifying common factors. Dimensionality reduction eg. PCA.

Assumptions:
Linearity, adequate sample size, absence of multicollinearity, and normality.

Tests Required:
KMO measure of sampling adequacy, Bartlett's test of sphericity, factor extraction method, and factor rotation method.

Apply the test and Validate

7

- Check for Sensitivity to assumptions
- Perform Sensitivity Analysis
- Test for Robustness

Report Results

8

- **Statistical Significance (Uncertainty):** p-value, confidence interval, and standard error.
- **Effect size:** Cohen's d, odds ratio, correlation coefficient
- **Sensitivity Analysis**
- Test for Robustness, bias, and fairness
- Provide future research direc