ISBN 978-1-960740-66-3

# SOCIOLOGY AND FAMILY STUDIES - FUNDAMENTAL AND APPLIED APPROACHES

*Book Chapter*

**STATISTICAL TEST SELECTION IN SOCIAL SCIENCES**

*October 17, 2025*
*doi: 10.5281/zenodo.14830086*

Scientific Knowledge Publisher (SciKnowPub), USA
info@sciknowpub.com

**BOOK CHAPTER**

# STATISTICAL TEST SELECTION IN SOCIAL SCIENCES

Nabeela Sulaiman[1*], Raheela Shahid[2]

[1]*Department of Education, Virtual University of Pakistan, Islamabad, Pakistan*
[2]*Department of Applied Psychology, Bahauddin Zakariya University, Multan, Pakistan*

***For Correspondence***
*nabeela.sulaiman@vu.edu.pk*

## Abstract

This chapter provides a comprehensive framework for selecting appropriate statistical techniques in social science research. The selection process involves four critical considerations: research objectives, the level of measurement for data, effect sizes, and assumptions about data distribution. By following this systematic approach, researchers can ensure their analytical choices align with their research goals and data characteristics.

## Keywords

Statistical Test, Social Sciences, Analytical Technique, Research Objectives, Assumptions, Data Distribution, Effect Size, Inferential Statistics

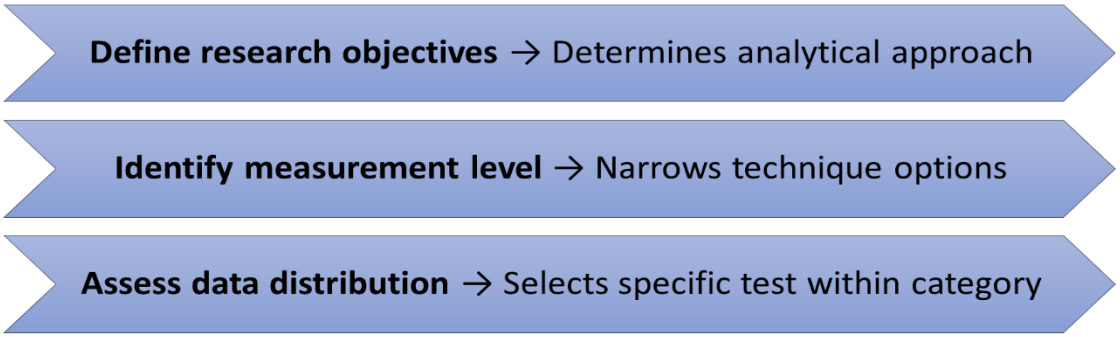
## Introduction

When selecting statistical and analytical techniques, four important considerations should be kept in mind. This systematic approach prevents common analytical errors and enhances the validity of research findings. The framework presented here has been validated across multiple social science disciplines and provides evidence-based decision criteria.

## Decision Framework Overview

The statistical test selection process follows a hierarchical decision tree (see Figure 1), this decision tree guides researchers through a systematic process, ensuring that analytical choices align with research objectives, data characteristics, and statistical assumptions. Each decision point builds upon

the previous one, creating a logical pathway to appropriate statistical techniques.




**Section 1: Research Objectives/Hypotheses**

Start by clearly defining the objectives of your analysis. Are you looking to describe patterns, make predictions, test hypotheses, or identify relationships? The goals of your research will guide your choice of statistical techniques. If your research involves testing specific hypotheses or making inferences about populations, statistical hypothesis tests are essential. Techniques such as t-tests, analysis of variance (ANOVA), chi-square tests, and z-tests help you assess the significance of observed differences and relationships in your data. In the realm of statistical analytics, the cornerstone of selecting appropriate techniques lies in defining your research objectives/hypotheses with precision. Your chosen statistical methods should be in direct alignment with these overarching goals as shown in Table 1.

**Table 1**. Quick Reference Decision Matrix

| Research Objective | Primary Question | Statistical Approach | Example Techniques |
|---|---|---|---|
| **Descriptive** | What are the characteristics? | Summarize and describe | Mean, median, frequency distributions |
| **Comparative** | Are groups different? | Test differences | t-tests, ANOVA, chi-square |
| **Relational** | How are variables related? | Examine associations | Correlation, regression |
| **Predictive** | What will happen? | Model outcomes | Regression, time-series |
| **Exploratory** | What patterns exist? | Discover structures | Factor analysis, clustering |

Here's a deeper exploration of this crucial consideration:

**A. Descriptive Analysis**

If your primary aim is to gain a better understanding of your data, describe patterns, and summarize key characteristics, then descriptive statistics should be your focal point. Key techniques which will provide essential insights into the data's central tendencies and dispersion include:

a) **Measures of central tendency**: Mean, median, mode

b) **Measures of variability**: Standard deviation, variance, range, interquartile range

c) **Distribution shape**: Skewness, kurtosis

d) **Visual summaries**: Histograms, box plots, frequency tables

**B. Comparative Analysis**

When your goal is to make comparisons among different sample groups or to differentiate between two sets of responses, you need to select the relevant statistical tests i.e., independent sample t-test, paired-sample t-test, one-way ANOVA, and two-way ANOVA. When comparing groups or conditions, your choice depends on:

**Number of groups**

a) **Two groups**: Independent samples t-test, paired t-test, Mann-Whitney U

b) **Multiple groups**: One-way ANOVA, Kruskal-Wallis test

c) **Multiple factors**: Two-way ANOVA, factorial designs

**Data characteristics**

a) **Normal distribution**: Parametric tests (t-tests, ANOVA)

b) **Non-normal distribution**: Non-parametric alternatives

c) **Small samples**: Consider exact tests or bootstrapping

**C. Relational Analysis**

Sometimes, the primary objective is to discover relationships or associations between variables. Correlation analysis, covariance analysis, and techniques like Pearson's correlation coefficient or Spearman's rank correlation coefficient can be valuable in elucidating such connections. To examine associations between variables:

**Correlation Analysis**

a) **Pearson's r**: Linear relationships between continuous variables

b) **Spearman's ρ**: Monotonic relationships, ordinal data

c) **Point-biserial**: Continuous and binary variables

d) **Phi coefficient**: Two binary variables

**Advanced Relational Techniques**

a) **Partial correlation**: Controlling for confounding variables

b) **Multiple regression**: Multiple predictors, one outcome

c) **Path analysis**: Complex theoretical models

**D. Predictive Modeling**

When the goal is to make informed predictions or forecasts, predictive analytics techniques come into play to estimate unknown values based on existing data patterns.

Regression analysis, time-series forecasting, and machine learning algorithms like linear regression, decision trees, and neural networks can be employed to build predictive models based on historical data patterns.

**Model Selection Criteria**

a) **Sample size**: Minimum 10-15 cases per predictor

b) **Variable types**: Determines model family (linear, logistic, etc.)

c) **Prediction accuracy requirements**: Influences complexity needs

**d) Interpretability needs**: Simple vs. complex models' trade-off

**E. Exploratory Data Analysis (EDA)**

EDA aims to uncover hidden insights, identify outliers, and visualize data structures. Tools like scatter plots, histograms, box plots, and correlation matrices are used to explore relationships and anomalies within the data. So,

**Systematic EDA Process**

EDA follows a structured approach:

a) **Initial data inspection**: Missing values, outliers, data quality
b) **Univariate exploration**: Individual variable distributions
c) **Bivariate relationships**: Pairwise associations
d) **Multivariate patterns**: Complex interactions and structures

**Essential EDA Tools**

a) **Graphical**: Scatter plots, correlation matrices, parallel coordinates
b) **Statistical**: Summary statistics, correlation coefficients
c) **Advanced**: Principal component analysis, cluster analysis

**F. Cluster Analysis**

If your aim is to group similar data points together or classify data into distinct categories, clustering techniques like k-means clustering and classification algorithms like logistic regression, support vector machines, and random forests are pertinent.

**G. Time-series Analysis**

For time-dependent data, time-series analysis methods such as auto-regressive integrated moving average (ARIMA) models and exponential smoothing can be employed to forecast future values and detect trends and seasonality.

**H. Causal Inference**

When research objectives involve establishing causality rather than merely association:

a) **Experimental designs**: Randomized controlled trials with manipulation of independent variables
b) **Quasi-experimental approaches**: Natural experiments, difference-in-differences analysis
c) **Observational causal methods**: Instrumental variables, regression discontinuity designs
d) **Longitudinal causal models**: Fixed effects, random effects, and growth curve models

**Key Considerations for Causal Claims**

a) **Temporal precedence**: Cause must precede effect
b) **Confounding control**: Account for alternative explanations
c) **Mechanism specification**: Theoretical justification for causal pathway
d) **Effect size interpretation**: Practical significance of causal effects

In summary, the choice of statistical techniques hinges on the clarity and precision with which you define your research objectives. Understanding the overarching purpose of your analysis, whether it's descriptive, predictive, hypothesis-driven, exploratory, or relationship-focused, lays the foundation for selecting the most suitable statistical methods to extract meaningful insights from your data.

**Section 2: Measurement Level of Data**

The measurement level of data is a critical factor in selecting the appropriate analysis technique in statistics. Determine the level of data you're working with: categorical (nominal or ordinal) or numerical (continuous or discrete). Different statistical methods are suited for different data measurement levels. Different statistical techniques are designed to work with specific types of data, and choosing the right one based on your data's characteristics is essential for accurate and meaningful results.

**Table 2.** Levels of Measurement in Research

| Level of Measurement | Definition | Key Properties | Examples (Measured Variables) |
|---|---|---|---|
| **Nominal** | Categories with **no inherent order** | Classification only | Gender, Religion, Ethnicity, Marital Status, Nationality, Blood Type, Yes/No Responses |
| **Ordinal** | Categories with **meaningful order**, but unequal intervals | Rank order only | Education Level, **Likert-type item** (e.g., *Strongly Disagree to Strongly Agree*), Socioeconomic Class, Military Rank, Pain Ratings |
| **Interval** | **Ordered** with **equal intervals**, but **no true zero** | Can measure **differences**, not ratios | Temperature (°C, °F), IQ Scores, SAT Scores, Time of Day, Calendar Years, **Composite Likert Scales** (if treated as continuous) |
| **Ratio** | Interval scale with a **true zero** point | Can measure **differences and ratios** | Height, Weight, Age, Income, Distance, Duration, Heart Rate, Reaction Time, Kelvin Temperature, Number of Children, Test Scores (out of 100) |

Note. Most temperature scales in practice have meaningful zeros (Kelvin, Rankine)

**Footnotes**

1. **Interval vs. Ratio**: Interval scales have equal units but **no true zero** (e.g., 0°C ≠ absence of temperature). Ratio scales have a **true zero**, allowing statements like "twice as much."

2. **Likert-Type Scale Examples**:

   a) **Single Item (Ordinal)**: I feel stressed at work.1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly

Agree

→ Treated as **ordinal** (no assumption of equal intervals).

b) **Composite Scale (Interval/Continuous Treatment):** Sum or average of responses across **5–10 related** items on stress, anxiety, etc.→ Often **treated as interval-level continuous data in statistical analyses** like regression, ANOVA, etc.

3. **Treatment of Likert Scales in Research**: While technically ordinal, composite Likert scales can be treated as interval if:

   a) They contain multiple items (typically 4+)
   b) The scale shows good reliability (e.g., Cronbach's a ≥ .70)
   c) Data is approximately normally distributed

4. **Time of Day**: (e.g., 3 PM vs. 6 PM) is **interval**, not ratio, because 0:00 is arbitrary and "twice as late" isn't meaningful.

5. **Discrete vs. Continuous**:

   a) **Nominal and ordinal** = discrete (categorical)

   b) **Interval and ratio** = usually continuous

Here's how measurement level or data type can guide your selection of analysis techniques:

**Nominal Data**

Nominal data is a categorical data type that represents categories or labels with no inherent order or ranking. In this type of data, each category is distinct and unrelated to the others, and there is no natural order or hierarchy among them. For example, gender (male/female/non-binary) treatment conditions (control/experimental) diagnostic categories (depression/anxiety/bipolar disorder) and ethnic groups.

## Appropriate Analysis Techniques

Analyzing nominal data, which consists of categories or labels with no inherent order or ranking, requires specific statistical techniques to draw meaningful insights. Here are some appropriate analysis techniques for nominal data:

### 1. Frequency Distribution

A basic step in analyzing nominal data is to create a frequency distribution. This summarizes the counts or frequencies of each category in your dataset. This simple technique provides an overview of the distribution of your nominal variable.

### 2. Mode Identification

In nominal data, the mode is the category that appears most frequently. Identifying the mode can provide insight into the most common category in your dataset.

### 3. Bar Charts

Bar charts or bar graphs are excellent visualizations for nominal data. They display the categories on the x-axis and the corresponding frequencies on the y-axis. Bar charts make it easy to compare the sizes of different categories.

### 4. Pie Charts

Pie charts are another graphical representation suitable for nominal data. They show the proportions of each category in a circular chart, making it visually clear how each category contributes to the whole.

## Statistical Tests for Association

### 1. Chi-Square Test of Independence

Test association between two nominal variables

- **a)** Use when: Expected frequencies ≥ 5 in all cells
- **b)** Effect size: Cramer's V for strength of association

The Chi-Square Test of Independence is a fundamental statistical test for analyzing relationships between two nominal variables. It helps determine if there is a statistically significant association between the variables. This test is particularly useful when you want to investigate if two nominal variables are dependent or independent of each other.

**2. Fisher's Exact Test**

Alternative to chi-square for small sample sizes (expected frequencies < 5). Fisher's Exact Test is an alternative to the Chi-Square Test, often used when dealing with small sample sizes. It's used to analyze contingency tables with nominal data.

**3. Contingency Tables (Cross-tabulations)**

Creating a contingency table, also known as a cross-tabulation or crosstab, is a standard approach when working with two nominal variables. It displays the joint frequencies of the categories, making it easier to assess relationships.

**Advanced Analytical Techniques**

**1. Multinomial Logistic Regression**

When you have multiple nominal independent variables and want to predict the probability of outcomes in a categorical dependent variable, multinomial logistic regression is appropriate. It extends logistic regression to handle more than two categories.

**2. Association Measures**

Measures like Cramer's V and Phi Coefficient quantify the strength of association between two nominal variables. These measures help assess the degree of dependence between variables beyond just determining significance.

**3. Mosaic Plots**

Mosaic plots are a graphical way to visualize the relationship between two or more categorical variables. They are particularly useful when analyzing associations among multiple nominal variables.

**4. Log-Linear Models**

Log-linear models are used when you have multiple nominal variables and want to model their relationships, considering interactions and associations among the variables.

**5. McNemar's Test**

Compare paired categorical responses (e.g., before/after treatment)

**Association Strength Measures**:

a) **Cramer's V**: Standardized measure (0 to 1) for any size contingency table

b) **Phi Coefficient**: For 2×2 tables specifically

c) **Odds Ratios**: Risk comparisons between categories

d) **Cohen's Guidelines**: Small (0.10), Medium (0.30), Large (0.50) for Cramer's V

**Ordinal Data**

Ordinal data has categories with a meaningful order or ranking but lacks equal intervals between them. Examples include Likert scales (1=strongly disagree to 5=strongly agree), education levels (high school < bachelor's < master's < doctorate), socioeconomic status (low < middle < high), and performance ratings (poor < fair < good < excellent. 1-star to 5-star). Ordinal data represents categories with a meaningful order or ranking, but the intervals between categories are not necessarily equal. Analyzing ordinal data requires specialized statistical techniques that account for the ordered nature of the categories.

## Appropriate Analysis Techniques

Here are some appropriate analysis techniques for ordinal data:

### Descriptive Statistics

**1. Frequency Distribution**

Start by creating a frequency distribution to understand the distribution of your ordinal data. This will provide an overview of the frequencies of each category.

**2. Bar Charts or Ordered Bar Charts**

Bar charts are useful for visualizing the distribution of ordinal data. In an ordered bar chart, the categories are arranged in a specific order to reflect their ordinal nature.

**3. Median and Quartiles**

Since ordinal data lacks equal intervals between categories, measures of central tendency like the median and quartiles (e.g., first quartile, third quartile) are often more appropriate than the mean. These statistics help describe the central tendency and spread of the data.

### Non-Parametric Tests

Non-parametric tests are useful for comparing ordinal data between groups or conditions because they do not assume a normal distribution. Common non-parametric tests for ordinal data include the Wilcoxon Rank-Sum Test (Mann-Whitney U Test) for comparing two groups and the Kruskal-Wallis Test for comparing more than two groups.

**1. Mann-Whitney U Test (Wilcoxon Rank-Sum)**: Compare two independent groups

- **a.** *Advantage*: No distribution assumptions
- **b.** *When to use*: Two groups, ordinal outcome variable

**2. Wilcoxon Signed-Rank Test**: Compare paired/related samples

**3. Kruskal-Wallis Test**: Compare multiple (3+) independent groups

**4, Friedman Test**: Compare multiple related groups (repeated measures design)

**Correlation and Regression**

1. **Ordinal Regression:**

   Ordinal regression, also known as ordered logistic regression or proportional odds model, is a specialized regression technique for modeling relationships between ordinal dependent variables and one or more independent variables. It accounts for the ordinal nature of the outcome variable and can provide valuable insights into predictors' effects.

2. **Spearman's Rank Correlation**:

   When you want to assess the strength and direction of a relationship between two ordinal variables, Spearman's rank correlation coefficient is appropriate. It measures the monotonic association between variables.

3. **Proportional Odds Model:**

   If you're conducting ordinal regression, the proportional odds model is a commonly used technique. It estimates the odds ratios associated with predictors and the cumulative log-odds of the ordinal outcome.

4. **Visual Inspection of Cumulative Probability Plots:**

   Cumulative probability plots can help visualize the cumulative probabilities of ordinal categories relative to predictor variables. This can provide insight into how ordinal data patterns vary with predictors.

5. **Polychronic and Polyserial Correlation:**

   When you have ordinal and continuous variables and want to assess correlations, polychronic correlation is suitable for ordinal-ordinal relationships, while polyserial correlation is used for ordinal-continuous relationships.

6. **Logistic Regression with Ordinal Dependent Variables:**

   In some cases, you might treat ordinal data as nominal data and use logistic regression. However, this approach doesn't fully account for the ordinal nature of the data, and ordinal regression is generally preferred.

7. **Categorical Data Analysis:**

   Techniques from categorical data analysis can be adapted for ordinal data analysis, particularly when dealing with multiple ordinal variables and exploring relationships between them.

**Important Note on Likert Scales**

**Likert Scale Controversy**: Many researchers treat Likert scales (1-5, 1-7 rating scales) as interval data when:

**a.** Multiple items are averaged into composite scores

**b.** Sample sizes are large (n > 30)

**c.** Distributions are approximately normal *Conservative approach*: Treat individual items as ordinal, composite scores as interval

**Interval Data**

True interval data features equal intervals between values but lacks a meaningful zero point. However, genuine interval data is extremely rare in the social sciences. Examples include Temperature scales: Celsius, Fahrenheit (but Kelvin and Rankine have true zeros), Calendar years: 2023, 2024 (arbitrary zero point), Standardized test scores: Some IQ tests, SAT scores (when scaled with an arbitrary zero).

Interval data represents categories with ordered values where the intervals between categories are equal and meaningful. Analyzing interval data typically involves techniques that account for the ordered nature of the data and its continuous nature. For instance, most variables traditionally labeled "interval" in social sciences (attitude scales, rating scales, test scores) are better treated as ratio data for analytical purposes, as they often have meaningful zero points or can be treated as such.

**Appropriate Analysis Techniques**

Here are some appropriate analysis techniques for interval data:

**1. Descriptive Statistics:**

Begin by calculating descriptive statistics, including the mean, median, and standard deviation. These measures provide insights into the central tendency, variability, and distribution of interval data.

**2. Histograms and Frequency Distributions:**

Visualize the distribution of interval data using histograms, frequency polygons, or stem-and-leaf plots. These graphs help you understand the shape and characteristics of the data.

**3. Box Plots:**

Box plots (box-and-whisker plots) are useful for displaying the distribution of interval data, highlighting the median, quartiles, and potential outliers.

**4. Parametric Tests (When Assumptions met):**

Interval data often follows a normal distribution. Parametric tests, such as t-tests for comparing means or analysis of variance (ANOVA) for comparing multiple groups, are suitable when assumptions of normality and homoscedasticity are met.

**5. Correlation Analysis:**

To assess relationships between interval variables, use Pearson's correlation coefficient. It measures the strength and direction of linear associations between two continuous variables.

**6. Regression Analysis:**

Linear regression is a powerful technique for modeling relationships between one or more independent variables and an interval-dependent variable. It helps you make predictions based on the data.

**7. ANOVA (Analysis of Variance):**

When you have multiple groups or factors, ANOVA can determine whether there are statistically significant differences in means. One-way ANOVA is used for one factor, while two-way or multi-way ANOVA involves multiple factors.

**8. Normality Tests:**

Use normality tests like the Shapiro-Wilk test or the Kolmogorov-Smirnov test to assess whether your interval data follows a normal distribution. This is important when applying parametric tests.

**9. Paired t-Test:**

If you want to compare two related samples, such as before-and-after measurements, the paired t-test assesses whether there is a significant difference in means.

**10. Analysis of Covariance (ANCOVA):**

ANCOVA combines analysis of variance (ANOVA) and regression. It allows you to compare means while controlling for the influence of one or more covariates.

**11. Survival Analysis:**

For time-to-event data, such as survival times or failure times, survival analysis techniques like Kaplan-Meier survival curves and Cox proportional hazards regression are appropriate.

**12. Multivariate Analysis:**

If you're dealing with multiple interval variables, multivariate techniques like principal component analysis (PCA) and factor analysis can help reduce dimensionality and identify patterns.

**13. Chi-Square Test for Normality:**

The chi-square test for normality assesses whether the distribution of interval data significantly deviates from a normal distribution.

Selecting the appropriate technique depends on your research objectives, the nature of your interval data, and the assumptions that hold for your dataset. Whether you're exploring relationships, making predictions, or

comparing groups, there are statistical methods tailored to address the characteristics of interval data.

**Note.** When assumptions are violated, then use Non-parametric alternatives**:** Treat data as ordinal data and use rank-based methods.

**Ratio Data**

Ratio data has ordered categories with equal intervals and a true zero point. Most continuous measures in social sciences are ratio data: age, income, and reaction time, number of correct responses, years of experience, height, weight, and frequency counts. Ratio data represents categories with ordered values, equal intervals between categories, and a true zero point. This data type allows for meaningful ratios between values and is considered the most versatile and informative type of data. Analyzing ratio data involves a wide range of statistical techniques, depending on your research objectives.

**Appropriate Analysis Techniques**

Here are some appropriate analysis techniques for ratio data:

**1. Descriptive Statistics:**

Start by calculating descriptive statistics, including the mean, median, standard deviation, and measures of spread and central tendency. These statistics provide insights into the central values and variability of ratio data.

**2. Histograms and Frequency Distributions:**

Visualize the distribution of ratio data using histograms, frequency polygons, or box plots. These graphs help you understand the shape and characteristics of the data.

**3. Scatterplots:**

Scatterplots are useful for visualizing relationships between two continuous ratio variables. They help identify trends, patterns, and outliers in the data.

**4. Outlier detection**:

Z-scores, box plot identification, leverage analysis

**5. Correlation Analysis:**

To quantify the strength and direction of linear associations between two ratio variables, use Pearson's correlation coefficient. It provides a measure of the linear relationship between variables.

**6. Chi-Square Test for Normality:**

This test assesses whether the distribution of ratio data significantly deviates from a normal distribution. It is useful for checking the assumption of normality before applying parametric tests.

**Inferential Statistics**

**Parametric tests (when assumptions met)**

**1. Regression Analysis:**

Linear regression is a powerful technique for modeling relationships between one or more independent variables and a continuous ratio-dependent variable. It allows you to make predictions based on the data.

**2. T-Tests:**

Use t-tests to compare means between two groups when you want to determine if the difference is statistically significant. Paired t-tests are appropriate for related samples.

**3. Analysis of Variance (ANOVA):**

When you have multiple groups or factors and want to assess whether there are statistically significant differences in means, ANOVA is a suitable technique. One-way ANOVA is used for one factor, while two-way or multi-way ANOVA involves multiple factors.

**4. Analysis of Covariance (ANCOVA):**

ANCOVA combines analysis of variance (ANOVA) and regression. It allows you to compare means while controlling for the influence of one or more covariates.

**5. Logistic Regression:**

When your dependent variable is binary (e.g., yes/no, success/failure) but you have continuous ratio predictors, logistic regression models the probability of the binary outcome.

**6. Log-linear Models:**

Log-linear models are used for modeling relationships between multiple categorical and continuous variables when dealing with ratio data.

**Non-Parametric tests (when assumptions not met)**

**1. Survival Analysis:**

For time-to-event data, such as survival times or failure times, survival analysis techniques like Kaplan-Meier survival curves and Cox proportional hazards regression are applicable.

**2. Cluster Analysis:**

Cluster analysis helps group similar cases based on similarity or dissimilarity measures calculated from ratio variables. Techniques like k-means clustering can uncover hidden patterns.

**3. Principal Component Analysis (PCA):**

PCA is a dimensionality reduction technique that is particularly useful when dealing with high-dimensional ratio data. It identifies patterns and reduces the number of variables while retaining most of the variance.

**Model Selection Criteria**

Choosing Among Options:

**a.** Sample size: Parametric tests more powerful with n ≥ 30 and normal distributions

**b.** Distribution shape: Non-parametric methods for severe skewness or outliers

**c.** Research purpose: Descriptive vs. inferential vs. predictive goals

**d.** Effect size requirements: Practical significance thresholds

Selecting the appropriate technique for ratio data analysis depends on your specific research goals, the nature of your data, and the assumptions that apply to your dataset. Whether you're exploring relationships, making predictions, comparing groups, or reducing dimensionality, there are statistical methods tailored to address the characteristics of ratio data.

**Section 3: Data Distribution and Assumptions**

Before selecting any statistical test, researchers must evaluate whether their data meets the technique's underlying assumptions. Violating assumptions can lead to incorrect conclusions, inflated Type I error rates, or reduced statistical power (see Table 3).

**Table 3.** Essential Assumption Categories

| Assumption Type | What It Means | How to Check | Violation Consequences |
|---|---|---|---|
| **Normality** | Data follows bell curve | Shapiro-Wilk test, Q-Q plots | Inflated Type I error |
| **Independence** | Observations unrelated | Study design consideration | Underestimated standard errors |
| **Homoscedasticity** | Equal variances | Levene's test, residual plots | Biased significance tests |
| **Linearity** | Straight-line relationships | Scatterplots, residual analysis | Poor model fit |

**Assumption Violation Severity Guide**

***Violation Impact Assessment:***

a. **Minor violations** (slight departures): Results generally robust, proceed with caution

b. **Moderate violations** (noticeable departures): Consider transformations or robust alternatives

c. **Severe violations** (major departures): Use non-parametric or robust methods mandatory.

**Parametric Tests for Normally Distributed Data**

In the realm of statistics, when dealing with data that follows a normal distribution, researchers often employ parametric tests. A normal distribution, also known as a Gaussian distribution or bell curve, is characterized by its symmetrical shape and is described by its mean and standard deviation. Parametric tests are statistical methods that make specific assumptions about the distribution of data, with the assumption of normality being a common one.

Parametric tests are particularly useful when working with normally distributed data because they rely on this assumption to make precise inferences about population parameters. Some widely used parametric tests for normally distributed data include the t-test and analysis of variance (ANOVA) for comparing means of two or more groups, as well as linear regression for modeling relationships between variables.

The t-test, for instance, can be employed to determine if there is a significant difference between the means of two groups, such as comparing the test scores of students in two different classes. ANOVA extends this concept to compare means among more than two groups, making it suitable for scenarios like testing the effectiveness of different teaching methods across multiple classes. Linear regression, on the other hand, allows

researchers to model and predict the relationship between variables, which can be valuable in various fields, from economics to psychology.

It's important to note that while parametric tests are powerful tools for normally distributed data, it's crucial to assess the assumption of normality before applying them, as deviations from a normal distribution can lead to inaccurate results. When data significantly deviates from normality, non-parametric tests, which make fewer assumptions about the data's distribution, may be more appropriate. Nonetheless, when working with normally distributed data, parametric tests provide researchers with valuable tools for hypothesis testing and statistical analysis. Thus, here is a quick guide:

**When to Use Parametric Tests**

a) **Sample size**: n ≥ 30 (Central Limit Theorem applies) OR

b) **Distribution**: Data approximately normal (Shapiro-Wilk p > 0.05)

c) **Homogeneity**: Equal group variances (Levene's test p > 0.05)

**Common Parametric Tests: (*t-tests*)**

a) **One-sample**: Compare sample mean to population value

b) **Independent samples**: Compare two group means

c) **Paired samples**: Compare related measurements

**ANOVA Family**

a) **One-way ANOVA**: Compare multiple group means

b) **Two-way ANOVA**: Two categorical predictors

c) **Repeated measures ANOVA**: Within-subjects designs

d) **ANCOVA**: Control for continuous covariates

**Regression Analysis**

a) **Simple linear**: One predictor, one outcome

b) **Multiple linear**: Multiple predictors, continuous outcome

c) **Polynomial**: Non-linear relationships

Certainly, let's explore "Non-Parametric Tests for Non-Normally Distributed Data" within the context of a "Non-Normal Distribution":

**Non-Parametric Tests for Non-Normally Distributed Data**

In statistics, there are situations where data does not conform to a normal distribution, meaning it doesn't display the characteristic bell-shaped curve and may exhibit skewed or irregular patterns. In such cases, researchers turn to non-parametric tests, which are statistical methods that do not rely on assumptions about the underlying distribution of data. Non-parametric tests are robust alternatives when dealing with non-normally distributed data, making them particularly valuable in various research scenarios.

Some common non-parametric tests include the Mann-Whitney U test, the Kruskal-Wallis test, and the Wilcoxon signed-rank test: Mann-Whitney U Test is employed when comparing two independent groups. It assesses whether there is a significant difference between the distributions of two samples. For example, it can be used to determine if there's a difference in the exam scores between two different teaching methods. When dealing with more than two independent groups, the Kruskal-Wallis test comes into play. It's an extension of the Mann-Whitney U test and helps determine if there are significant differences among the groups. An example could be comparing the effectiveness of three different medications on pain relief. Explaining furthermore, Wilcoxon Signed-Rank Test is designed for paired or related samples. It assesses whether there's a significant difference between two related sets of data. For instance, it can be used to evaluate whether there's a change in the performance of students before and after a specific intervention.

Non-parametric tests are valuable because they don't require assumptions about the data's distribution and are more robust in the presence of outliers or highly skewed data. However, they may have less statistical power compared to their parametric counterparts when the data does, in fact, follow a normal distribution. Therefore, the choice between parametric and non-parametric tests should depend on the nature of the data and the research question, with non-parametric tests being the preferred option when normality assumptions are not met. Thus, here is a quick approach to deciding about tests:

**Decision criteria and power considerations**

**When to Choose Non-Parametric Tests**

a) **Distribution violations**: Severely skewed, multiple peaks, heavy tails

b) **Small samples**: n < 30 with unknown distribution

c) **Ordinal data**: Ranked but not interval-level measurements

d) **Outliers present**: Extreme values affecting means

**Power Considerations**

a) **When normality holds**: Non-parametric tests are ~95% as powerful as parametric tests

b) **When normality violated**: Non-parametric may be MORE powerful

c) **Sample size impact**: Power differences decrease with larger samples

d) **Rule of thumb**: With n > 100, power differences become minimal

**Common Non-Parametric Tests**

**Two-Group Comparisons:**

a) **Mann-Whitney U**: Independent samples (alternative to t-test)

b) **Wilcoxon signed-rank**: Paired samples (alternative to paired t-test)

***Multiple-Group Comparisons:***

a) **Kruskal-Wallis**: Independent groups (alternative to one-way ANOVA)
b) **Friedman test**: Related groups (alternative to repeated measures ANOVA)

***Post-hoc Tests (Important follow-up):***

a) **Dunn's test**: Follow-up to Kruskal-Wallis

b) **Wilcoxon pairwise**: Multiple comparisons with correction

**Assumption Violation Solutions**

**Data Transformation Options:**

a) **Log transformation**: Right-skewed data → log(x+1)

b) **Square root**: Moderate skew → √x

c) **Reciprocal**: Severe right skew → 1/x

d) **Box-Cox**: Optimal transformation determination

**Robust Statistical Approaches**

a) **Bootstrapping**: Generate sampling distribution empirically

b) **Trimmed means**: Remove extreme percentiles

c) **Robust regression**: Resistant to outliers

d) **Permutation tests**: Distribution-free significance testing

**Alternative Strategies (When Transformations Fail)**

a. **Outlier removal:** Only with theoretical justification
b. **Separate analysis by subgroups**: When heterogeneity is meaningful
c. **Bayesian approaches:** Less dependent on distributional assumptions
d. **Effect size focus:** Emphasize practical over statistical significance

## Section 4: Sample Size and Power Considerations

Statistical power represents the probability of detecting a true effect when it exists. Adequate power (typically ≥ 0.80) is essential for meaningful research conclusions.

### Power Analysis Components

a) **Effect size**: Magnitude of difference/relationship you want to detect

b) **Alpha level**: Type I error rate (typically 0.05)

c) **Sample size**: Number of observations

d) **Statistical power**: Probability of detecting true effects (1 - β)

### Sample Size Guidelines

**Basic Comparative Tests:**

a) **t-tests**: Minimum 15-20 per group for medium effects, Cohen's d = 0.5.

b) **ANOVA**: 20-30 per group recommended, for f = 0.25.

c) **Chi-square**: Expected frequencies ≥ 5 in all cells, minimum total N = 20.

### Regression Analysis

a) **Simple regression**: Minimum 50 + 8m (m = predictors), for $R^2 \geq 0.13$

b) **Multiple regression**: 104 + m for testing $R^2$; 50 + 8m for individual predictors

c) **Logistic regression**: Minimum 10 events per predictor variable, preferably 15-20.

### Power Analysis Tools, Recommended Software

a. **G*Power (free):** Comprehensive power analysis for most common tests
b. **R packages:** pwr, WebPower for advanced calculations
c. **Online calculators**: Sample size calculators for basic designs

d. **Post-hoc power:** Generally discouraged - focus on effect sizes instead

In summary, selecting the right statistical analytics techniques and sample size involves a thoughtful assessment of your research goals, data characteristics, assumptions, and practical considerations. Making informed choices based on these considerations will help ensure the validity and relevance of your statistical analysis.

## Section 5: Effect Sizes and Practical Significance

Statistical significance tells us if an effect exists; effect size tells us if it matters practically.

### Cohen's Conventional Effect Size Guidelines:

**For Mean Differences (Cohen's d):**

a. Small effect: d = 0.20 (means differ by 0.2 standard deviations)

b. Medium effect: d = 0.50 (means differ by 0.5 standard deviations)

c. Large effect: d = 0.80 (means differ by 0.8 standard deviations)

**For Correlations (Pearson's r):**

a. Small effect: r = 0.10 (1% of variance explained)

b. Medium effect: r = 0.30 (9% of variance explained)

c. Large effect: r = 0.50 (25% of variance explained)

**For Categorical Associations (Cramer's V):**

a. Small effect: V = 0.10

b. Medium effect: V = 0.30

c. Large effect: V = 0.50

### Reporting Standards:

a. Always report effect sizes alongside p-values

b. Include confidence intervals for effect sizes when possible
c. Interpret practical significance in context of research domain
d. Consider minimal important difference based on theory or prior research

**Context-Dependent Interpretation:**

a. Educational interventions: d = 0.25 may be practically significant
b. Medical treatments: Small effects (d = 0.20) can be life-saving
c. Psychological interventions: Medium to large effects (d ≥ 0.50) typically expected
d. Survey research: Small correlations (r = 0.10-0.20) may represent important associations

**Section 6: Practical Decision-Making Framework**

Follow this systematic approach for statistical test selection:

**Step 1: Clarify Research Question**

a) **Descriptive**: "What are the characteristics?"
b) **Comparative**: "Are groups different?"
c) **Relational**: "How are variables associated?"
d) **Predictive**: "What determines the outcome?"

**Step 2: Identify Variables**

a) **Dependent variable**: Outcome of interest
b) **Independent variables**: Predictors or grouping factors
c) **Measurement levels**: Nominal, ordinal, interval, ratio
d) **Number of variables**: Univariate, bivariate, multivariate

**Step 3: Check Assumptions**

a) **Sample size adequacy**: Power analysis

b) **Distribution normality**: Statistical tests and plots

c) **Independence**: Study design evaluation

d) **Homogeneity**: Equal variances across groups

**Step 4: Select Technique**

a) **Primary choice**: Based on research question and variable types

b) **Assumption violations**: Consider alternatives or transformations

c) **Effect sizes**: Plan for practical significance interpretation

**Step 5: Interpret Results**

a) **Statistical significance**: p-values and confidence intervals

b) **Practical significance**: Effect sizes and real-world importance

c) **Limitations**: Assumption violations and their impacts

**Step 6: Report and Communicate**

a) **Transparent reporting**: Document all analytical decisions and their rationale

b) **Complete results**: Include descriptive statistics, effect sizes, and confidence intervals

c) **Limitation acknowledgment**: Discuss assumption violations and their potential impact

d) **Practical implications**: Translate statistical findings into actionable insights

## The Key Takeaways in Selecting Analytical Techniques

## Summary: Evidence-Based Test Selection

### Core Decision Principles

a) **Objective-First Approach**: Define research questions before selecting analyses
b) **Data-Driven Decisions**: Let measurement level guide technique options
c) **Assumption Awareness**: Check and address violations appropriately
d) **Power Planning**: Ensure adequate sample sizes for meaningful results
e) **Effect Size Focus**: Report practical alongside statistical significance

### Quality Assurance Checklist

- Research objective clearly defined
- Variables and measurement levels identified
- Sample size adequate for planned analyses
- Assumptions checked and violations addressed
- Effect sizes calculated and interpreted
- Results interpreted in practical context

### Common Pitfalls to Avoid

a) Testing multiple analyses until finding significance (p-hacking)
b) Ignoring assumption violations
c) Focusing solely on p-values without considering effect sizes
d) Using inappropriate techniques for data types
e) Failing to report confidence intervals
f) Over-interpreting non-significant results

### Troubleshooting Guide

a) **When assumptions are violated**: Consider transformations → robust methods → non-parametric alternatives

b) **When sample size is small**: Use exact tests → bootstrap methods → report limitations

c) **When effect sizes are small**: Increase sample size → consider practical significance → report confidence intervals

**When in Doubt**

a) **Consult Figure 1**: Follow the decision tree systematically
b) **Use conservative approaches:** Non-parametric tests when assumptions are questionable
c) **Seek statistical consultation:** Complex designs warrant expert guidance
d) **Document decisions:** Justify your analytical choices in research reports

The decision framework presented in Figure 1 provides a visual roadmap for navigating the process of selecting statistical tests. Begin by defining your research objectives, identifying your data characteristics, and assessing distributional assumptions. Then, follow the flowchart to determine the appropriate analytical techniques.

This systematic framework ensures that statistical technique selection is driven by research objectives and data characteristics rather than convenience or familiarity. By following this evidence-based approach, researchers can enhance the validity and impact of their findings while avoiding common analytical pitfalls.

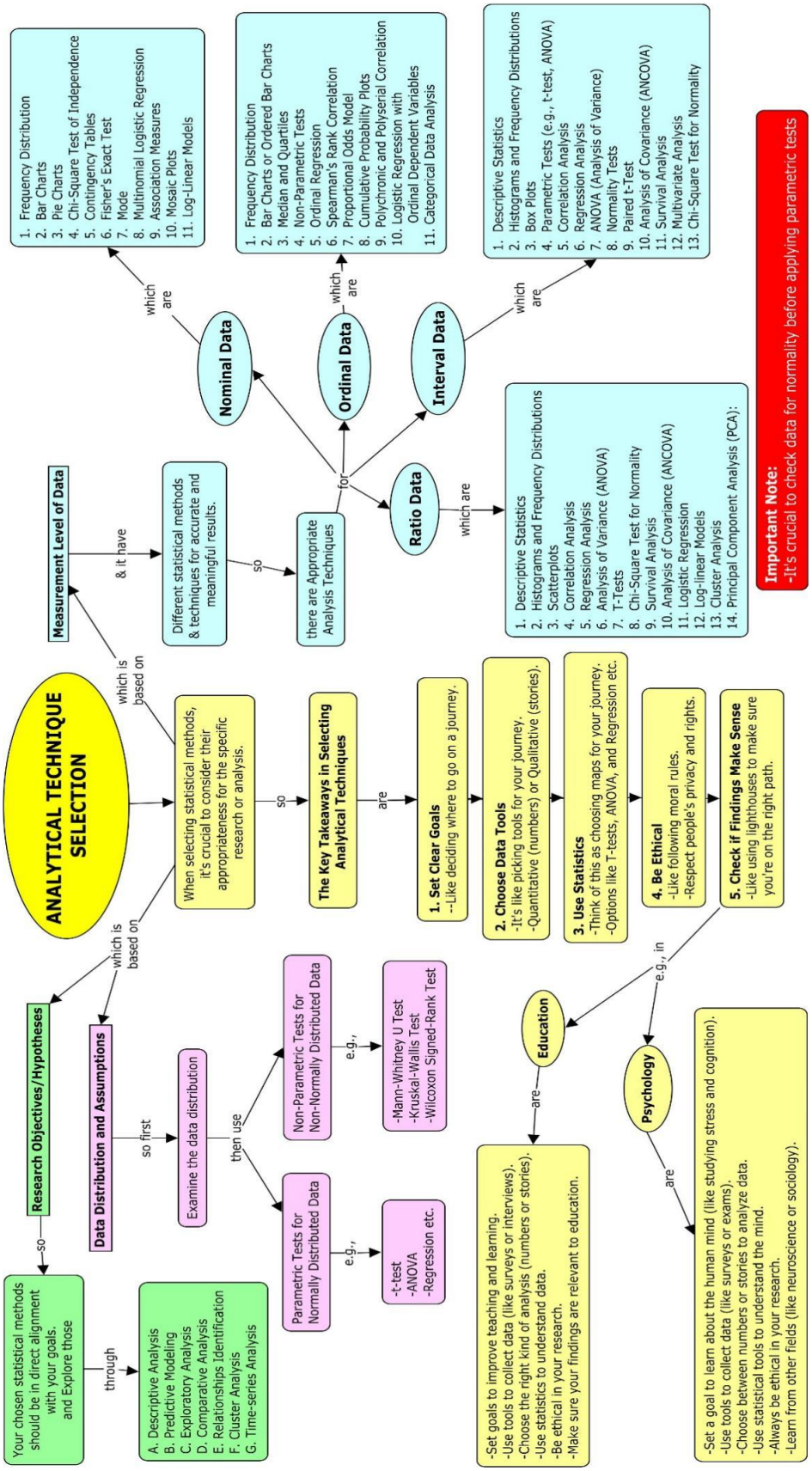



**Figure 1.** Concept Map of Considerations in Analytical Technique Selection

## Appendix A: Quick Reference Tables

Most Common Statistical Tests by Research Scenario

| Scenario | Variables | Recommended Test | Alternative |
|---|---|---|---|
| **Compare 2 group means** | 1 continuous DV, 1 binary IV | Independent t-test | Mann-Whitney U |
| **Compare paired means** | 1 continuous DV, paired data | Paired t-test | Wilcoxon signed-rank |
| **Compare 3+ group means** | 1 continuous DV, 1 categorical IV | One-way ANOVA | Kruskal-Wallis |
| **Relate 2 continuous variables** | 2 continuous variables | Pearson correlation | Spearman correlation |
| **Predict continuous outcome** | Multiple predictors, 1 continuous DV | Multiple regression | Robust regression |
| **Compare proportions** | 2 categorical variables | Chi-square test | Fisher's exact test |

**Acknowledgements**

We acknowledge the contribution of AI in reflecting the concepts in broader terms from other available sources on internet.

**Conflicts of interest**

The author does not have conflict of interest.