Choosing the Right Statistical Test

[Your Name/Affiliation]

[Current Date: 6/23/2025]

Why Does Test Selection Matter?

Choosing the correct statistical test is paramount for the integrity and validity of your research. The implications of selecting an inappropriate test can be far-reaching, leading to erroneous conclusions and potentially misleading results.

- **Drawing Accurate Conclusions:** The right test ensures that the inferences drawn from your data are statistically sound and reflect the true relationships or differences within your study.
- **Avoiding Misleading Results:** An incorrect test can produce false positives or false negatives, leading to misinterpretations of your findings and potentially flawed decision-making.
- **Ensuring Research Validity:** Proper test selection upholds the scientific rigor of your study, making your results credible and reproducible.
- **Ethical Considerations:** In fields like medicine or public policy, incorrect statistical analysis can have serious ethical implications, potentially leading to harmful interventions or misallocation of resources.

Key Questions Before Choosing a Test

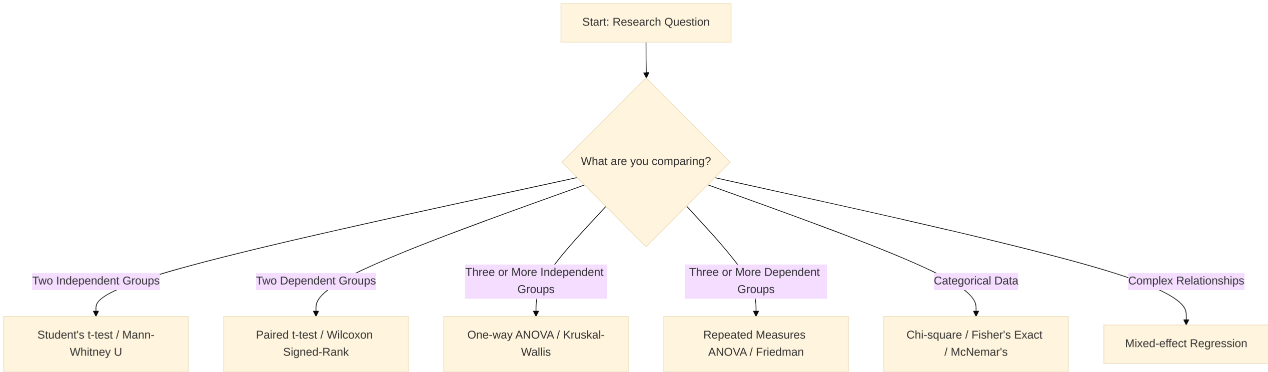
Before diving into statistical analysis, it's crucial to ask a series of fundamental questions about your research and data. These questions will guide you toward the most appropriate statistical test.

- **What is your research question?**

    - Are you looking for a difference between groups (e.g., Is there a difference in test scores between two teaching methods)?
    - Are you investigating a relationship between variables (e.g., Is there a correlation between study hours and exam performance)?
    - Are you examining an effect over time (e.g., Does a new drug reduce symptoms over several weeks)?
- **What type of data do you have?**
    - **Nominal:** Categorical data without order (e.g., gender, blood type).
    - **Ordinal:** Categorical data with a meaningful order but unequal intervals (e.g., Likert scales, education levels).
    - **Interval:** Numerical data with equal intervals but no true zero point (e.g., temperature in Celsius/Fahrenheit).
    - **Ratio:** Numerical data with equal intervals and a true zero point (e.g., height, weight, income).
- **How many groups/variables are you comparing?**
    - Are you comparing two groups, three or more groups, or looking at relationships between multiple variables?
    - Are the groups independent (e.g., different participants in each group) or paired/dependent (e.g., same participants measured at different times)?
- **Are your data normally distributed?**
    - This question helps determine whether to use **parametric** tests (which assume data follow a specific distribution, often normal) or **non-parametric** tests (which do not make such assumptions).
- **Are your observations independent?**
    - **Independent samples:** Observations in one group do not influence observations in another group.
    - **Dependent (paired) samples:** Observations are related, such as measurements taken from the same individuals at different times or matched pairs.

The Statistical Test Selection Flowchart (Overview)

This flowchart provides a high-level roadmap for selecting the appropriate statistical test. It guides you through the key decision points based on your research question and data characteristics. Subsequent slides will delve into each comparison type in more detail.



- **Visual:** The diagram above represents a simplified flowchart. In a real presentation, this would be a visually engaging graphic that acts as a navigation tool for the audience, showing the overall structure of the decision-making process.

Comparing Two Independent Groups

Research Question: Is there a significant difference between two unrelated groups?

Parametric Option: Student's t-test

- **Assumptions:**
  - Interval/Ratio data.
  - Independence of observations.
  - Approximate normality of the dependent variable within each group.

- Homogeneity of variances (variances of the dependent variable are equal across groups).

- **When to Use:** Comparing the means of two independent groups. For example, comparing the average test scores of students who received different teaching methods.

If Assumptions Violated (especially normality/homogeneity):

Alternative: Mann-Whitney U test

- **Assumptions:**
  - Ordinal/Interval/Ratio data.
  - Independence of observations.
  - The distributions of the two groups are similar in shape (though not necessarily normal).
- **When to Use:** Comparing the medians (or distributions) of two independent groups when the data are not normally distributed, or when the data are ordinal. For example, comparing satisfaction levels (ordinal) between two different customer service approaches.

Comparing Two Dependent (Paired) Groups

Research Question: Is there a significant difference between two related measurements (e.g., pre/post, matched pairs)?

Parametric Option: Paired t-test

- **Assumptions:**
  - Interval/Ratio data.

  - Independence of pairs (each pair is independent of other pairs).
  - Normality of the *differences* between the paired measurements.
- **When to Use:** Comparing the means of two paired samples. For example, comparing a patient's blood pressure before and after a treatment, or comparing the performance of matched individuals under two different conditions.

If Assumptions Violated (especially normality of differences):

Alternative: Wilcoxon signed-rank test

- **Assumptions:**
  - Ordinal/Interval/Ratio data.
  - Independence of pairs.
  - Symmetry of the distribution of the differences (though less strict than normality).
- **When to Use:** Comparing the medians (or distributions) of two paired samples when the data are not normally distributed, or when the data are ordinal. For example, assessing changes in pain scores (ordinal) before and after an intervention.

Comparing Three or More Independent Groups

Research Question: Is there a significant difference among three or more unrelated groups?

Parametric Option: One-way ANOVA (Analysis of Variance)

- **Assumptions:**
  - Interval/Ratio data.
  - Independence of observations.
  - Approximate normality of the dependent variable within each group.
  - Homogeneity of variances (variances of the dependent variable are equal across groups).
- **When to Use:** Comparing the means of three or more independent groups. For example, comparing the effectiveness of three different fertilizers on crop yield.

If Assumptions Violated (especially normality/homogeneity):

Alternative: Kruskal-Wallis H test

- **Assumptions:**
  - Ordinal/Interval/Ratio data.
  - Independence of observations.
  - The distributions of the groups are similar in shape (though not necessarily normal).
- **When to Use:** Comparing the medians (or distributions) of three or more independent groups when the data are not normally distributed, or when the data are ordinal. For example, comparing the ranking of customer satisfaction across three different product versions.

Comparing Three or More Dependent (Repeated Measures) Groups

Research Question: Is there a significant difference across three or more related measurements (e.g., multiple time points, multiple conditions on the same subjects)?

Parametric Option: Repeated Measures ANOVA

- **Assumptions:**
  - Interval/Ratio data.
  - Independence of subjects (each subject is independent of other subjects).
  - Normality of the dependent variable within each condition.
  - Sphericity (the variances of the differences between all combinations of related groups are equal). While less critical with modern software that can adjust for violations, it's an important assumption to be aware of.
- **When to Use:** Comparing means across three or more related samples. For example, tracking changes in a patient's pain level at baseline, 1 month, and 3 months after treatment.

If Assumptions Violated (especially normality/sphericity):

Alternative: Friedman test

- **Assumptions:**
  - Ordinal/Interval/Ratio data.
  - Independence of subjects.
- **When to Use:** Comparing medians (or distributions) across three or more related samples when the data are not normally distributed, or when the data are ordinal. For example, ranking the preference of three different product designs by the same group of users.

More Complex/Flexible Option: Mixed-effect Regression

Mixed-effect models (also known as multilevel models or hierarchical linear models) are powerful statistical tools used when data have a hierarchical or clustered structure. This often occurs in repeated measures designs where observations are nested within individuals, or in studies where individuals are nested within groups (e.g., students within classrooms, patients within hospitals).

- **When to Use:**
    - When you have hierarchical or clustered data, where observations are not independent.
    - For repeated measures designs, especially when there are missing data points or unequal time intervals.
    - To model individual trajectories and group effects simultaneously.
    - When you need to account for both fixed effects (variables whose effects are constant across individuals/groups) and random effects (variables whose effects vary across individuals/groups).
- **Assumptions:**
    - The assumptions for mixed-effect models are more flexible and depend on the specific model chosen. Generally, they include:
        - Linearity (for linear mixed models).
        - Independence of residuals (after accounting for the random effects).
        - Appropriate distribution for the outcome variable (e.g., normal for continuous outcomes, binomial for binary outcomes).
- **Note:** This is a more advanced statistical topic. The key takeaway is its utility for handling complex data structures and repeated measures where traditional ANOVA methods might fall short or be less efficient. It allows for more nuanced modeling of variability at different levels of the data hierarchy.

Analyzing Categorical Data

Research Question: Are observed frequencies/proportions different from expected, or is there an association between two categorical variables?

For Independent Categorical Data:

Chi-square test (χ^2)

- **Assumptions:**
  - Independent observations.
  - Expected cell counts should be at least 5 in most cells (or a reasonable proportion, e.g., no more than 20% of cells have expected counts less than 5, and no cell has an expected count less than 1).
- **When to Use:**
  - **Test of Association:** To determine if there is a significant association between two categorical variables (e.g., Is there an association between gender and political party affiliation?).
  - **Goodness-of-Fit Test:** To determine if observed frequencies for a single categorical variable differ significantly from expected frequencies (e.g., Does the distribution of colors in a bag of candies match the manufacturer's stated proportions?).

Fisher's exact test

- **Assumptions:**
  - Independent observations.
- **When to Use:** Preferred over the Chi-square test when dealing with small sample sizes, especially in 2x2 contingency tables, where the expected cell counts assumption for Chi-square is violated. It calculates the exact probability of observing the given data (or more extreme data) under the null hypothesis.

For Dependent Categorical Data:

McNemar's test

- **Assumptions:**
    - Paired nominal data (e.g., before-after measurements on the same subjects).
- **When to Use:** Comparing proportions in paired samples when the outcome is binary (nominal). For example, assessing whether there is a significant change in the proportion of people who answer a question correctly before and after an intervention, or comparing the agreement between two raters on a binary outcome.

Before You Run the Test: Data Exploration is Key!

Before applying any statistical test, thorough data exploration is a critical first step. This process helps you understand the characteristics of your data, identify potential issues, and verify assumptions required by parametric tests. Visual inspection is often more informative and intuitive than relying solely on statistical tests for assumptions.

- **Visualizing Data:**
    - **Histograms:** To visualize the distribution of a single continuous variable and check for normality, skewness, or multimodality.
    - **Box Plots:** To display the distribution of a continuous variable across different categories, identify outliers, and compare central tendency and spread between groups.
    - **Scatter Plots:** To examine the relationship between two continuous variables, identify patterns, and detect outliers or non-linear relationships.
- **Checking for Outliers:** Outliers can significantly influence statistical test results. Visualizations like box plots and scatter plots are excellent for identifying them.

Further investigation is needed to determine if outliers are data entry errors, measurement errors, or true extreme values.

- **Assessing Normality:** Many parametric tests assume that the data (or residuals) are normally distributed. While visual inspection (histograms, Q-Q plots) is crucial, formal tests can also be used:
    - **Q-Q Plots (Quantile-Quantile Plots):** A graphical method to compare the distribution of your data against a theoretical normal distribution. Points should fall approximately along a straight line.
    - **Shapiro-Wilk Test / Kolmogorov-Smirnov Test:** Statistical tests for normality. *Caution:* For large sample sizes, these tests can be statistically significant even with minor deviations from normality, which may not practically impact the robustness of the parametric test. Visual inspection and understanding the context of your data are often more important.
- **Assessing Homogeneity of Variances:** Some parametric tests (e.g., independent samples t-test, ANOVA) assume that the variances of the dependent variable are equal across groups.
    - **Levene's Test:** A statistical test to assess the equality of variances. A non-significant p-value suggests that the variances are equal.
    - **Visual Inspection:** Box plots can also give a quick visual indication of whether the spread of data is similar across groups.

A Checklist for Test Selection

This checklist provides a structured approach to selecting the appropriate statistical test. It can also serve as a useful handout for your audience.

Checklist Item	Your Answer
Research Question	What are you trying to find out? (e.g., difference, relationship, effect)
Number of Variables/Groups	How many are involved? (e.g., two, three+, paired, independent)
Variable Type(s)	What kind of data do you have? (e.g., Nominal, Ordinal, Interval, Ratio)
Independence of Observations	Are your samples paired or independent?
Distribution	Is your data normally distributed or non-normal?
Assumptions Met?	Have you checked the assumptions for your chosen test? (e.g., Yes/No)
Chosen Test	What is the primary statistical test you will use?
Alternative Test (if needed)	If assumptions are not met, what is your backup test?

Thank You

Thank you for your attention. I hope this presentation has provided you with a clear framework for selecting the right statistical test for your research.

Q&A

I would be happy to answer any questions you may have.