Data Analysis Using SPSS

Parametric and Non-parametric Tests

By

Dr.Anura Karunarathne

Senior Lecturer,

Department of Accountancy

University of Kelaniya

Choosing the Correct Statistical Test

Questions to ask yourself:

1. What **scale of measurement** has been used?
2. Are you testing **differences** between groups or **associations** between variables? **Purpose**
3. Are the samples **independent or related** (same person tested twice or paired)?
4. **How many groups** are involved?
5. Nature of Distribution of Your Data-**Normality?**

How to Use SPSS: Choosing the Appropriate Statistical Analysis Technique

Choosing the Right Statistic

- Matching research design to appropriate analysis
- Useful to follow a step-by-step process with 3 basic questions to be answered
  - What type of research question are you asking?
  - What type & number of variables do you want to analyze?
  - What type of data do you have and what characteristics does it have?

Step 1:What question type do you want to answer?

- Descriptive
  - What is the typical blood pressure for people in a given population?
- Correlational/Predictive
  - Is there a relationship between blood pressure and SES?
  - Can neighborhood predict blood pressure?
- Group Differences/Cause & Effect
  - Is the typical blood pressure higher in one residential area compared to another?

Overview of Common Statistical Analyses

- Descriptive Analysis
  - Frequency, percentiles, central tendency, standard scores
- Correlational Analysis
  - Correlation
  - Regression
- Analyzing Differences Between Groups
  - t-tests
  - One-Way ANOVA
  - Factorial ANOVA
  - MANOVA

Step 2: Identify & define your variables

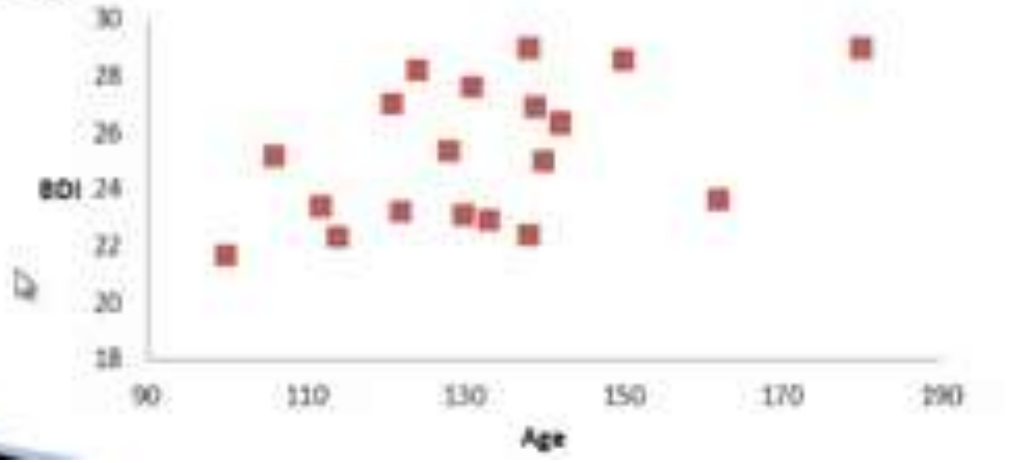
- Independent vs. dependent
- Number of each
- Operational definition of each variable
- Normal/expected range of scores/levels of each variable
  - Independent variables
  - Dependent variables

Step 3:Identify the nature of the variables

- Level of measurement for each variable
  - Nominal/categorical
  - Ordinal
  - Interval/ratio
- Additional information
  - Interval/ratio
    - Normally distributed?
    - Appropriate range of scores?
  - Nominal/categorical
    - Are the groups equal/balanced?
    - Are some of the categories empty?

Step 4: Draw a diagram of your design

- Summarize key points in a diagram
- Identify type of question & Identify variables as specifically as possible
- Example 1:Is there a relationship between blood pressure and body weight?
  - Body weight-continuous; body mass in pounds from 100-250
  - Blood pressure-continuous; pressure in mmHg from 100-220



Step 4: Draw a diagram of your design

- Example 2: Do people with BMI values below 25 have lower SBP than people with BMI above 25?
  - BMI value-independent, categorical (two groups/levels): BMI $\leq$ 25/BMI $\geq$ 25.1
  - SBP-dependent; continuous: mmHg range from 100-220
  -

	BMI $\leq$ 25	BMI $\geq$ 25.1
Mean Systolic BP (mmHg)		

Step 4:Draw a diagram of your design

- Example 3: Is the effect of sex on SBP different for people with BMI values below 25 than people with BMI above 25?
  - Sex-independent, categorical: males/females
  - BMI value-independent, categorical (two groups/levels): BMI $\leq$ 25/BMI $\geq$ 25.1
  - SBP-dependent; continuous: mmHg range from 100-220
  -

	BMI $\leq$ 25	BMI $\geq$ 25.1
Mean SBP for males		
Mean SBP for females		

Step 5: Determine need for parametric or non-parametric test

- Does your data meet the assumptions of parametric testing?
- What if it doesn't?
  - Use parametric testing anyway
  - Transform the data
  - Use a non-parametric technique
    - Pearson Correlation/Spearman Correlation or Chi-square for Independence
    - Independent t-test/Mann-Whitney U Test
    - Dependent t-test/Wilcoxon Signed Rank test
    - One-Way ANOVA/Kruskal-Wallis Test

Step 6: Make the final decision

- Make determinations about your variables
- Make sure you meet all the assumptions
- Are there other approaches that could be taken?
- What approach have other studies with similar designs used?

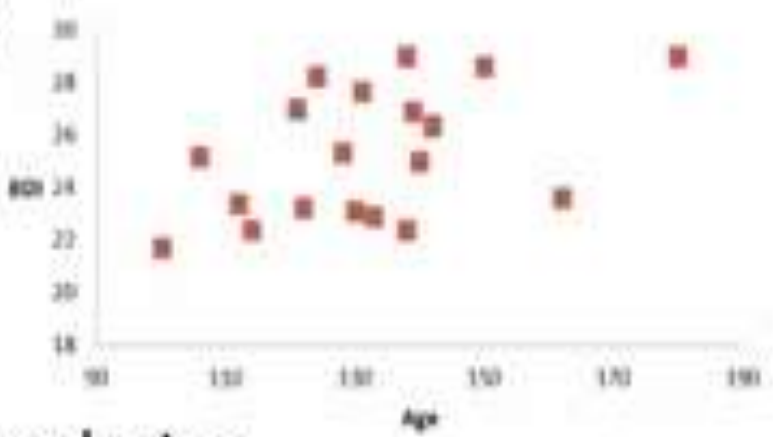
Example 1

- Research Question: What is the relationship between gender and having a diagnosis of clinical depression?
- What you have:
  - One categorical independent variable (IV); sex: male/female
  - One categorical dependent variable (DV); diagnosis of depression: yes/no (i.e. # of people in each category)

	Males	Females
Has depression		
Does not have depression		

Example 2

- Research Question: Is there a relationship between age and depression index? Does depression increase with age?
- What you have:
  - Two continuous variables (age, score on a depression index (BDI))



  - Technique: Pearson Correlation

Example 4

- Research Question: Will 10 weeks of exercise training reduce the BDI score?
- What you have:
  - One categorical IV: (pre-test/post-test)
  - One continuous DV (BDI score)

	Pre-test	Post-test
Mean BDI score		

  - Technique: Dependent/paired-sample t-test
  - Non-parametric alternative: Wilcoxon Signed Rank Test

Example 5

- Is there are difference in BDI score for people under 30, 31-49 and 50 years and over?
- What you have:
  - One categorical IV with two or more groups/levels (age: under 30, 31-49, 50+)
  - One continuous DV (BDI score)

	<30 years	31-49 years	50+ years
Mean BDI score			

  - Technique: One-Way ANOVA

Example 6

- Research Question: What is the effect of age on BDI scores for males and females?
- What you have
  - Two categorical IVs (sex: male/female; age: <30, 31-49, >50)
  - One continuous DV (BDI score)

	<30 years	31-49 years	> 50 years
Mean male BDI score			
Mean Female BDI score			

  - Technique: Factorial (Two-Way) ANOVA
  - Non-parametric Alternative: None

Example 7

- RQ: Which of two therapy interventions is more effective in reducing BDI score across three time periods (pre-treatment, post-treatment, 3 months post-treatment)?
- What you have:
  - One between groups IV: type of intervention (exercise, meditation)
  - One within-groups IV: (3 measurement points)
  - One continuous DV (BDI score)

	Time 1	Time 2	Time 3
Mean BDI with exercise			
Mean BDI with meditation			

  - Technique: Factorial Repeated Measures (Mixed Between-Within or Split-Plot) ANOVA

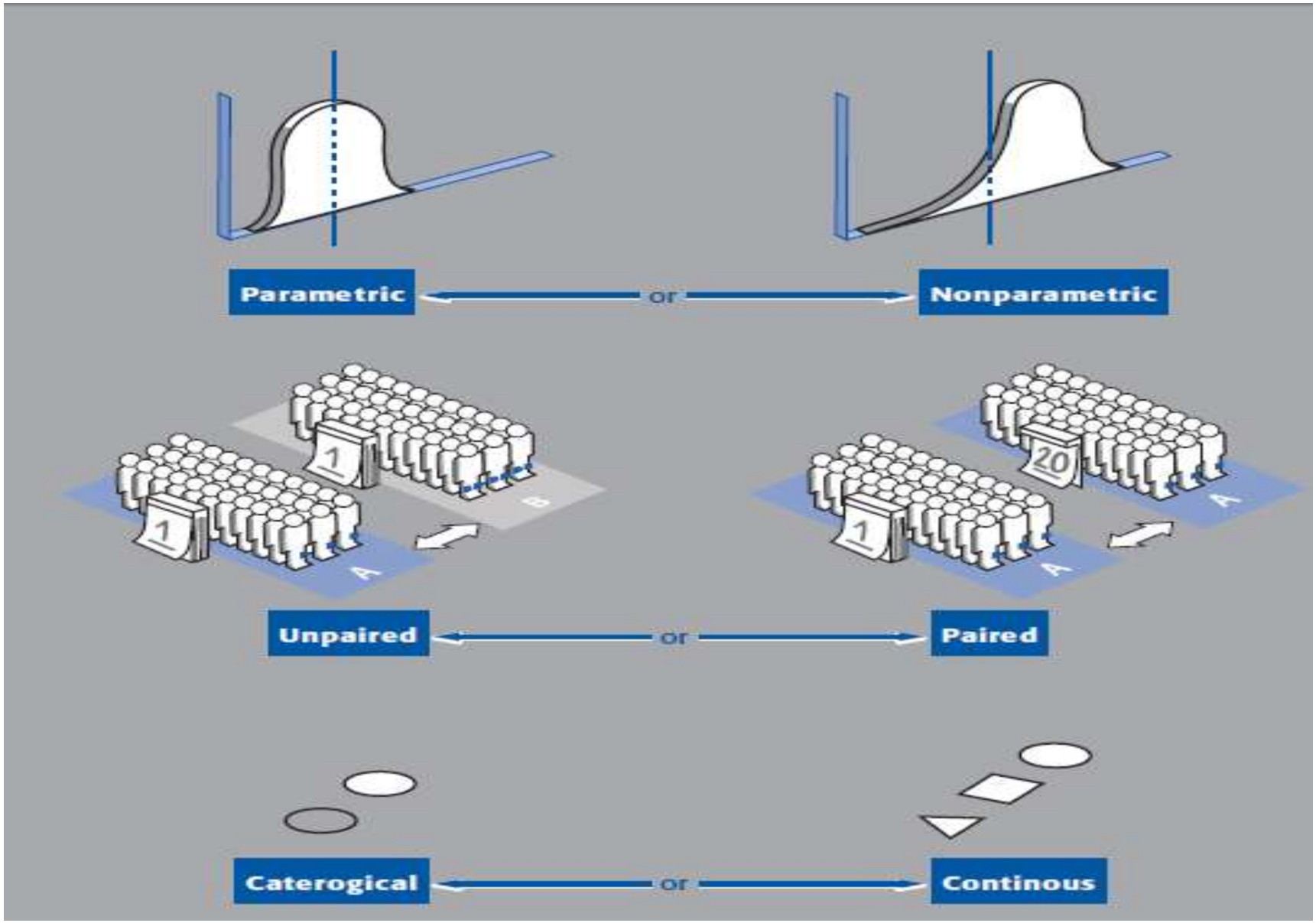
Example 8

- RQ: Do males have better overall ratings of psychological health (depression, anxiety, perceived stress) than females?
- What you have
  - One categorical IV (sex: male/female)
  - Two or more continuous DVs (depression, anxiety, perceived stress measures)

	Males	Females
Mean Anxiety score		
Mean Depression score		
Mean Perceived Stress score		

  - Technique: MANOVA
  - Non-parametric alternative: None

How to analyze your data



Summary.....

Type of Analysis by data type

Analysis	**Uni-variate Data**	**Bivariate Data**
Exploratory Data analysis/ Descriptive Analysis	**1. Presenting Frequencies** •Table •Graphical forms **2. Measuring Location** /Measures of Central Tendency •Mean •Median •Mode **3. Measuring Dispersion** •Range and Inter quartile Range •Standard Deviation/Variation **4. Measuring change** •Index Numbers	**1. Presenting Frequencies** •Cross Tabulation •Graphical forms •Scatter Diagrams •Stem Plots

Type of Analysis by data type…

Confirmatory Data Analysis/Inferential Statistics	1. Estimation From Samples •**Confidence Intervals (P)** **2. Forecasting** •**Time Series Analysis**	**1. Measuring Association** •Pearson's Correlation Coefficient-r (P) •Spearman's Rank Correlation Coefficient-r_k (NP) **2. Measuring Differences** Chi Squared (X^2) Test(NP) **3. Student t- test(P)**

Note:

P = Parametric Technique

NP = Non- Parametric Technique

Scale of Measurement

Type of Statistic		Scale of Measurement
Nonparametric	–	Nominal
Nonparametric	–	Ordinal
Nonparametric	Parametric	Approximately Interval
–	Parametric	Interval
–	Parametric	Ratio

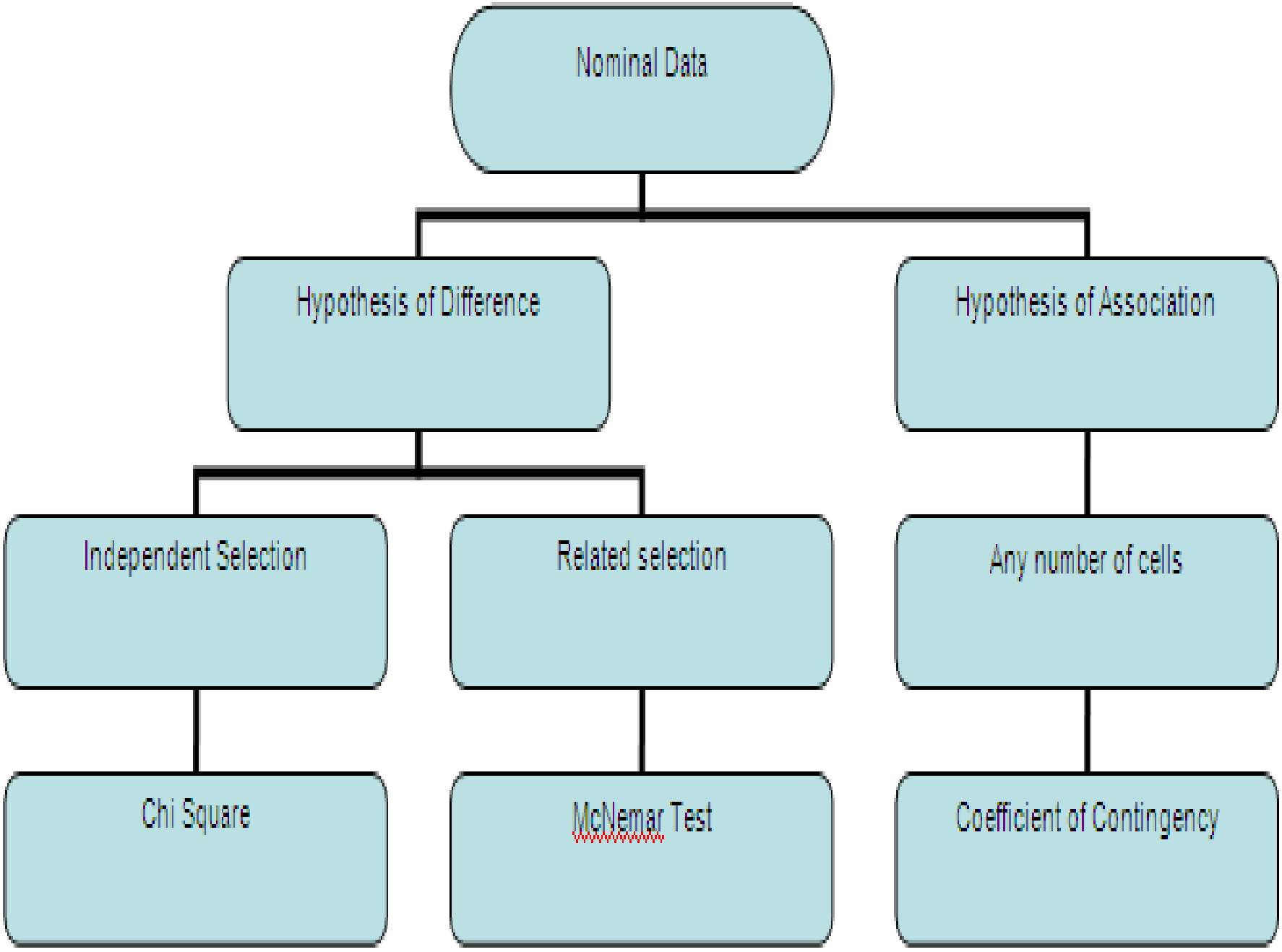
Nonparametric statistics

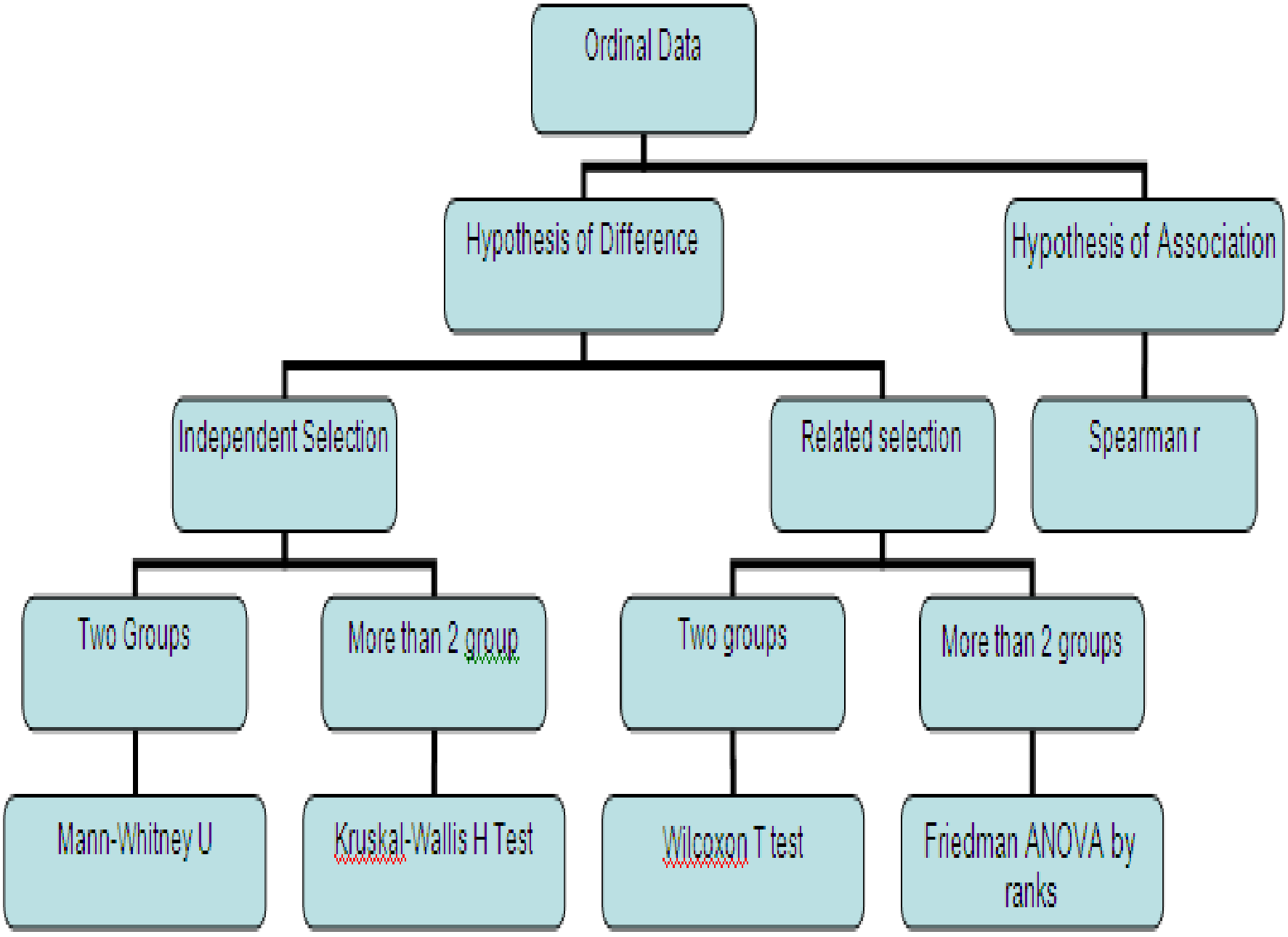
Nominal Data	Ordinal Data
Chi-Square Goodness-of-Fit Test	Mann-Whitney *U* Test
Chi-Square Test of Independence	Wilcoxon *T* Test
McNemar Test	Kruskal-Wallis *H* Test
	Friedman ANOVA by Ranks
	Spearman's r_s

Parametric statistics

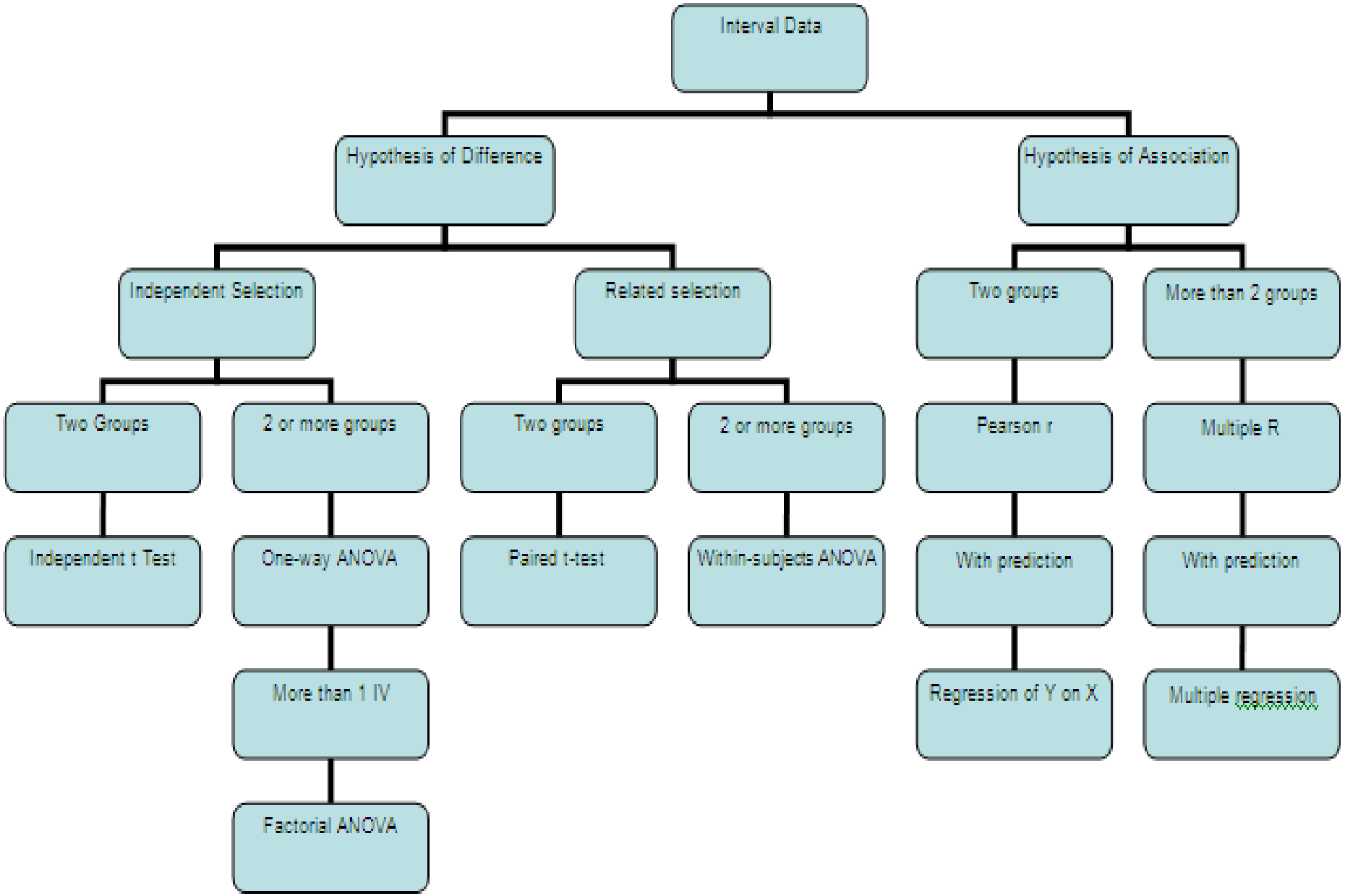
Parametric statistics are used when our data are measured on **approximately interval, interval, or ratio** scales of measurement.

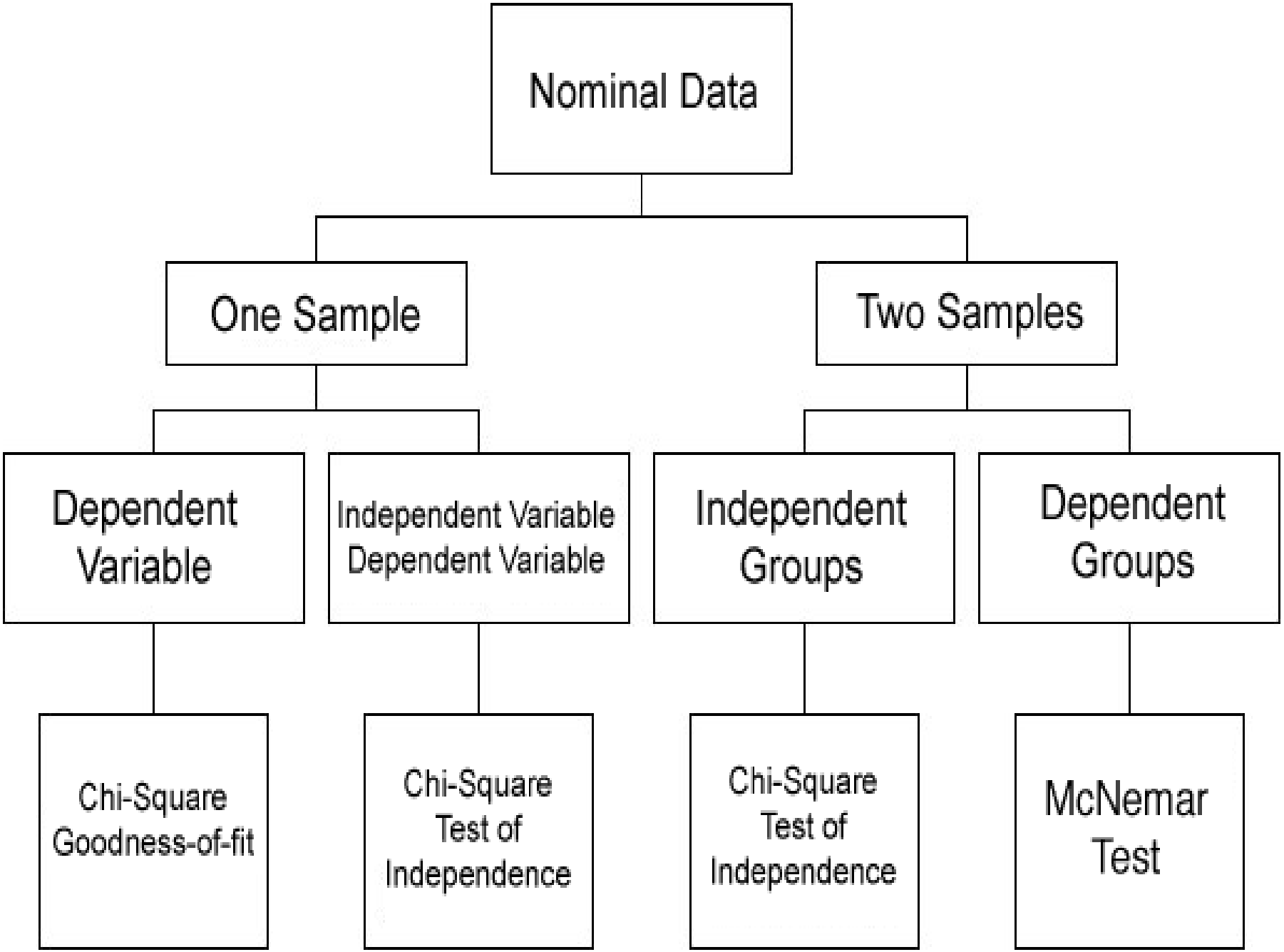
Parametric Statistics
Z Test
One-Sample *t* Test
t Test for Dependent Means
t Test for Independent Means
One-Way Between Groups ANOVA
Simple Repeated Measures ANOVA
Factorial ANOVA
Mixed Factorial ANOVA
Pearson's *r*
Bivariate Regression
Multiple Regression

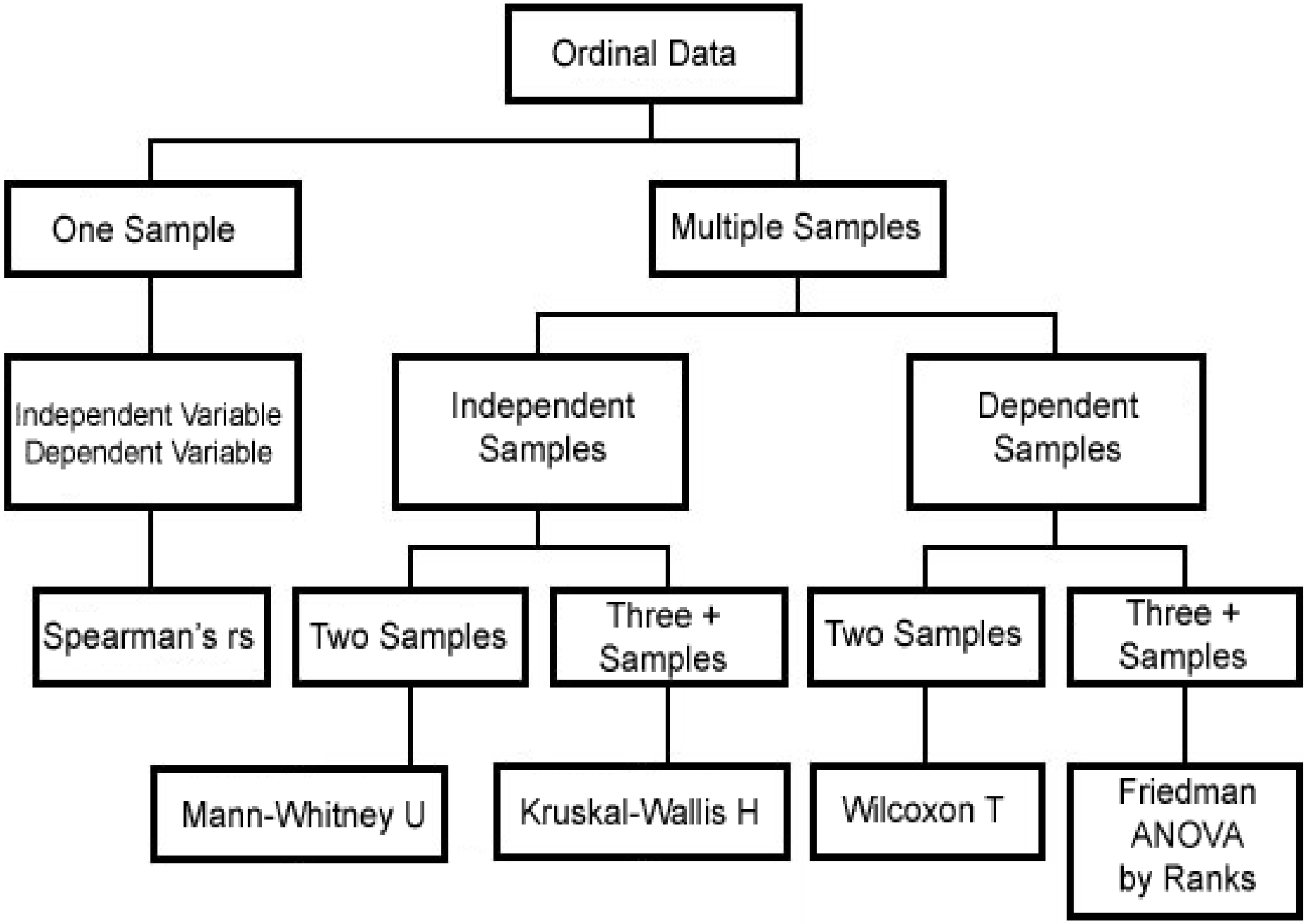

Nominal Data
Hypothesis of Difference
Hypothesis of Association
Independent Selection
Related selection
Any number of cells
Chi Square
McNemar Test
Coefficient of Contingency

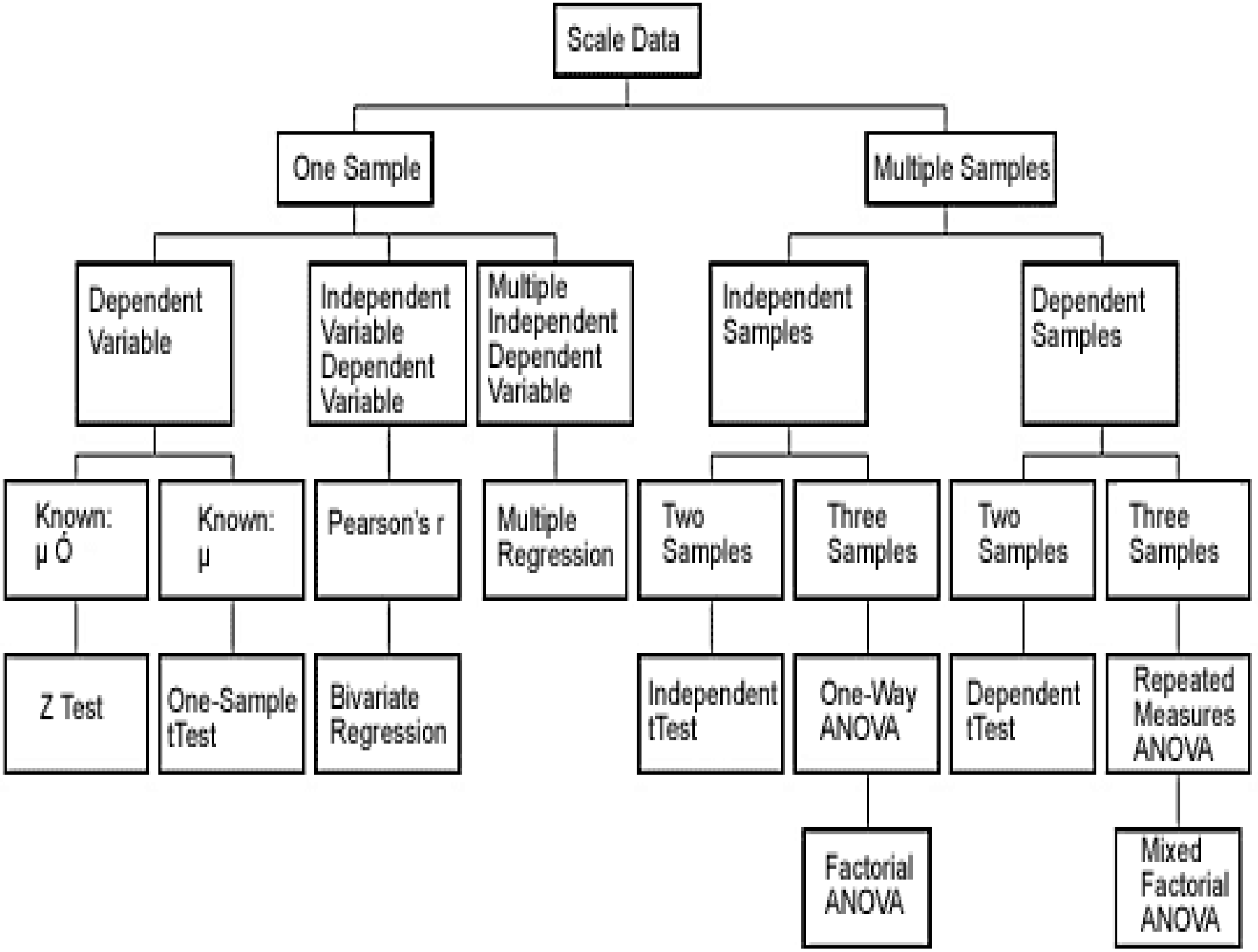

Ordinal Data
Hypothesis of Difference
Hypothesis of Association
Independent Selection
Related selection
Spearman r
Two Groups
More than 2 group
Two groups
More than 2 groups
Mann-Whitney U
Kruskal-Wallis H Test
Wilcoxon T test
Friedman ANOVA by ranks

Interval/ratio data:




Nominal Data
One Sample
Two Samples
Dependent Variable
Independent Variable Dependent Variable
Independent Groups
Dependent Groups
Chi-Square Goodness-of-fit
Chi-Square Test of Independence
Chi-Square Test of Independence
McNemar Test


Ordinal Data
One Sample
Multiple Samples
Independent Variable
Dependent Variable
Independent Samples
Dependent Samples
Spearman's rs
Two Samples
Three + Samples
Two Samples
Three + Samples
Mann-Whitney U
Kruskal-Wallis H
Wilcoxon T
Friedman ANOVA by Ranks


Scale Data
One Sample
Multiple Samples
Dependent Variable
Independent Variable Dependent Variable
Multiple Independent Dependent Variable
Independent Samples
Dependent Samples
Known: μ Ó
Known: μ
Pearson's r
Multiple Regression
Two Samples
Three Samples
Two Samples
Three Samples
Z Test
One-Sample tTest
Bivariate Regression
Independent tTest
One-Way ANOVA
Dependent tTest
Repeated Measures ANOVA
Factorial ANOVA
Mixed Factorial ANOVA