How to choose a statistical test

Francisco J. Candido dos Reis
DGO-FMRP
University of São Paulo

Choosing the right test

- One of the most common queries in stats support is ‘Which analysis should I use’
- There are several steps to help the student decide
- When a student is explaining their project, these are the questions you need answers for

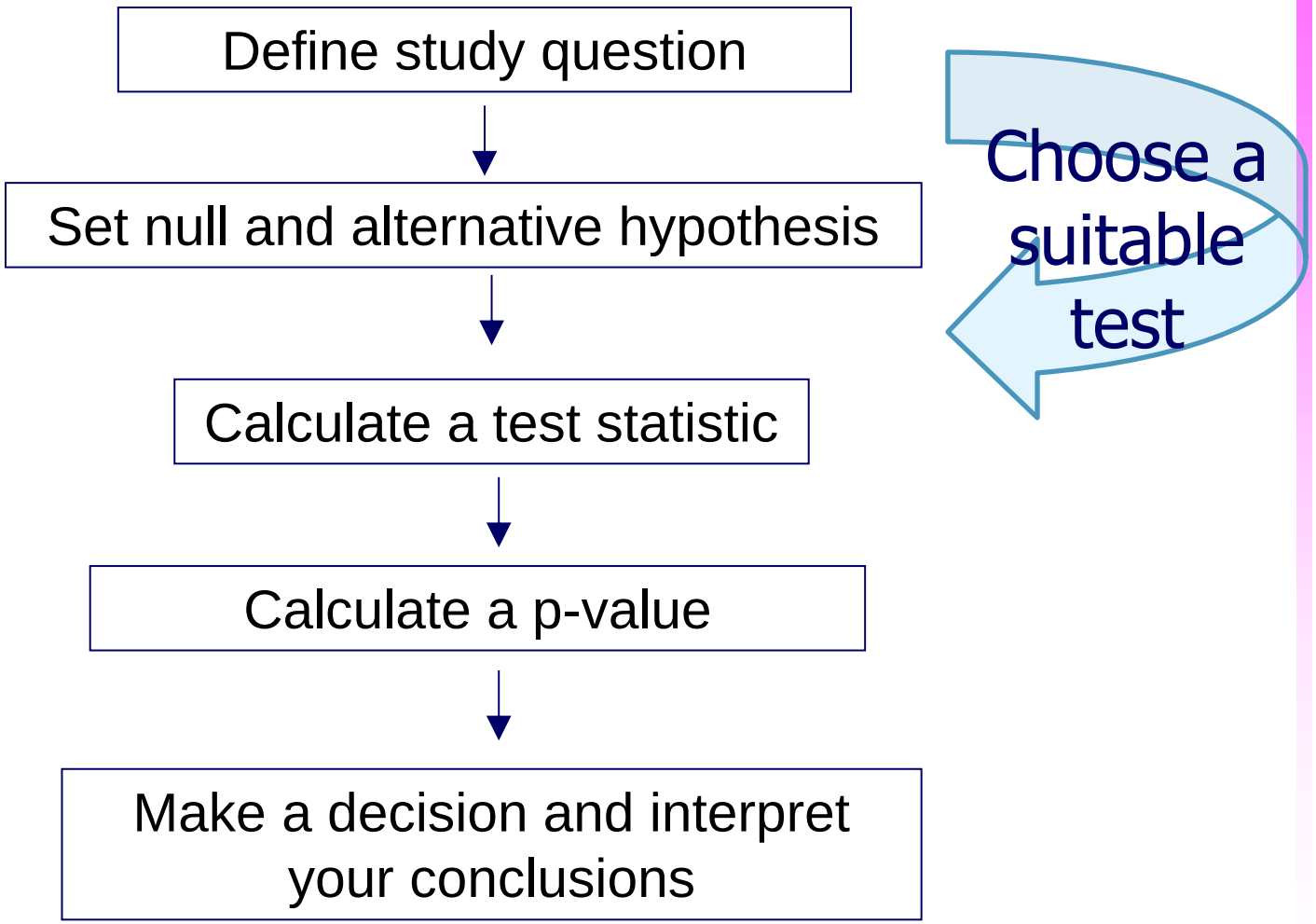
Choosing the right test

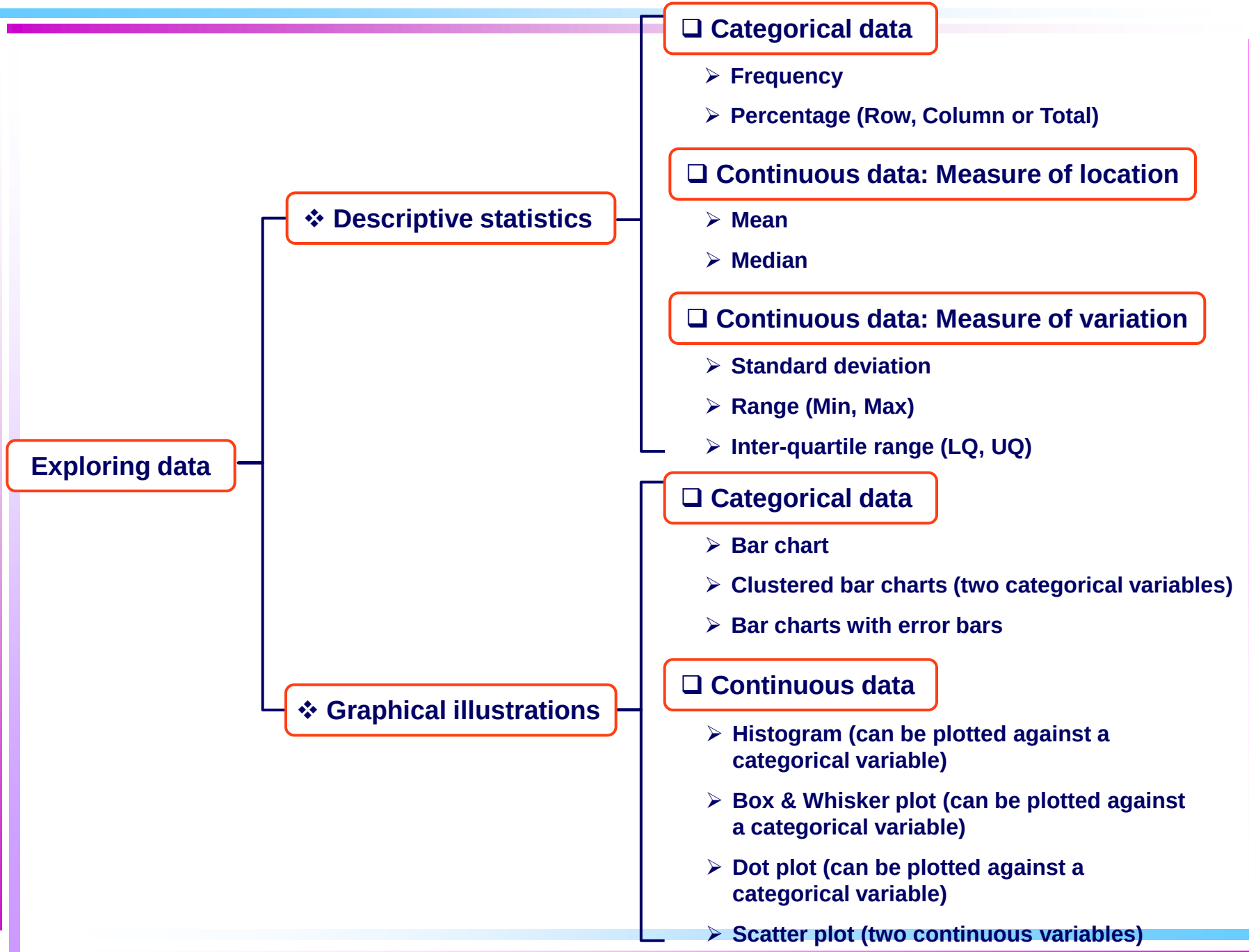
1) A clearly defined research question
2) What is the dependent variable and what type of variable is it?
3) How many independent variables are there and what data types are they?
4) Are you interested in comparing means or investigating relationships?
5) Do you have repeated measurements of the same variable for each subject?

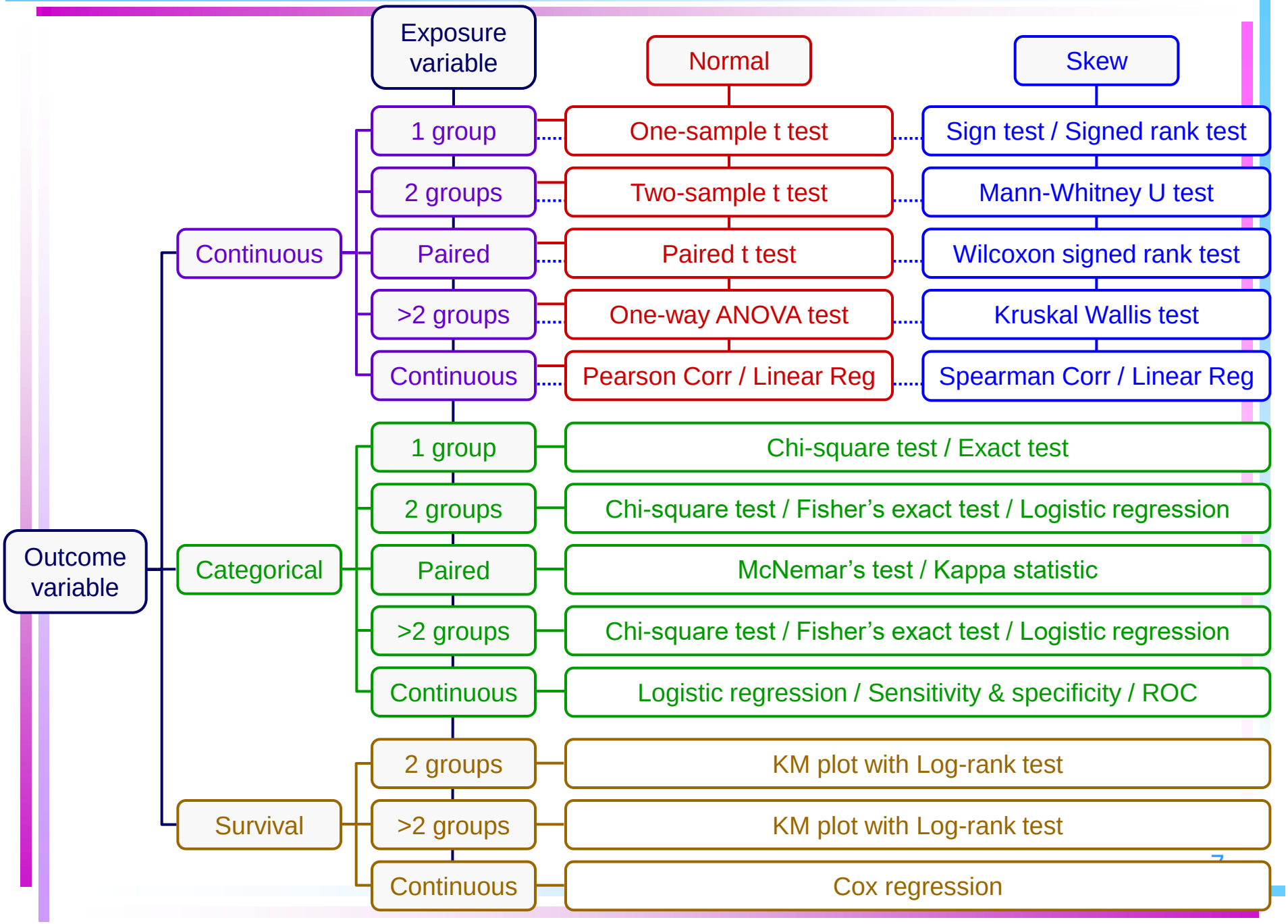
Research question

- Clear questions with measurable quantities
- Which variables will help answer these questions
- Think about what test is needed before carrying out a study so that the right type of variables are collected

Steps to undertaking a Hypothesis test




Exploring data
❖ Descriptive statistics
❑ Categorical data
➢ Frequency
➢ Percentage (Row, Column or Total)
❑ Continuous data: Measure of location
➢ Mean
➢ Median
❑ Continuous data: Measure of variation
➢ Standard deviation
➢ Range (Min, Max)
➢ Inter-quartile range (LQ, UQ)
❖ Graphical illustrations
❑ Categorical data
➢ Bar chart
➢ Clustered bar charts (two categorical variables)
➢ Bar charts with error bars
❑ Continuous data
➢ Histogram (can be plotted against a categorical variable)
➢ Box & Whisker plot (can be plotted against a categorical variable)
➢ Dot plot (can be plotted against a categorical variable)
➢ Scatter plot (two continuous variables)

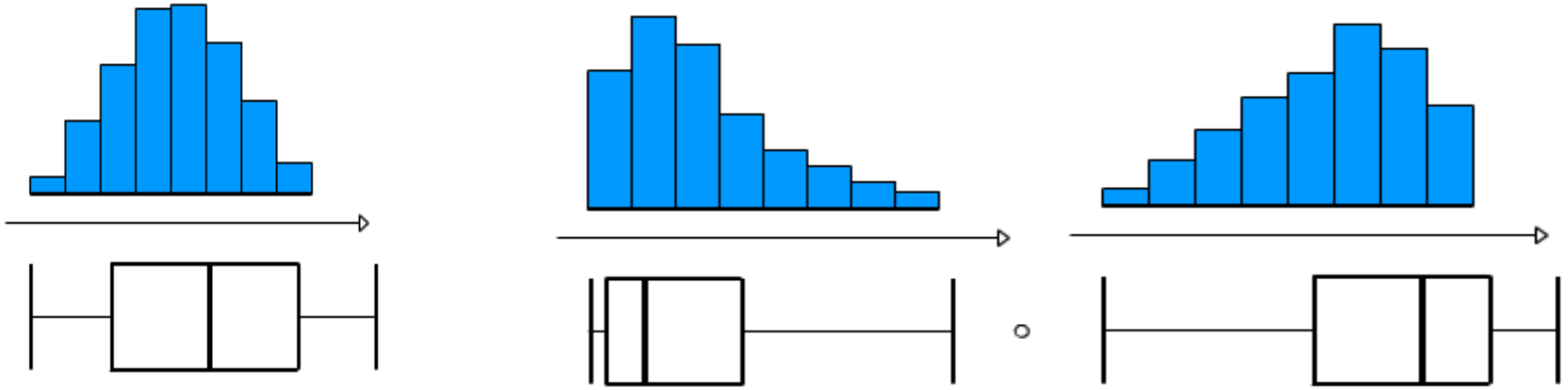

Exposure variable
Normal
Skew
Outcome variable
Continuous
1 group
One-sample t test
Sign test / Signed rank test
2 groups
Two-sample t test
Mann-Whitney U test
Paired
Paired t test
Wilcoxon signed rank test
>2 groups
One-way ANOVA test
Kruskal Wallis test
Continuous
Pearson Corr / Linear Reg
Spearman Corr / Linear Reg
Categorical
1 group
Chi-square test / Exact test
2 groups
Chi-square test / Fisher's exact test / Logistic regression
Paired
McNemar's test / Kappa statistic
>2 groups
Chi-square test / Fisher's exact test / Logistic regression
Continuous
Logistic regression / Sensitivity & specificity / ROC
Survival
2 groups
KM plot with Log-rank test
>2 groups
KM plot with Log-rank test
Continuous
Cox regression

Assessing Normality

Charts can be used to **informally** assess whether data is:

Normally distributed

Or….Skewed



The mean and median are very different for skewed data.

Chi-squared test statistic

- The chi-squared test is used when we want to see if two categorical variables are related
- The test statistic for the Chi-squared test uses the sum of the squared differences between each pair of observed (O) and expected values (E)

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

T-tests
Paired or Independent (Unpaired) Data?

T-tests are used to compare two population means

- **Paired data:** same individuals studied at two different times or under two conditions PAIRED T-TEST
- **Independent:** data collected from two separate groups INDEPENDENT SAMPLES T-TEST

Assumptions in t-Tests

- **Normality:** Plot histograms
  - One plot of the paired differences for any paired data
  - Two (One for each group) for independent samples
  - Don't have to be perfect, just roughly symmetric
- **Equal Population variances:** Compare sample standard deviations
  - As a rough estimate, one should be no more than twice the other
  - Do an F-test to formally test for differences
- However the *t*-test is very robust to violations of the assumptions of Normality and equal variances, particularly for moderate (i.e. >30) and larger sample sizes

What if the assumptions are not met?

- There are alternative tests which do not have these assumptions

Test	Check	Equivalent non-parametric test
Independent t-test	Histograms of data by group	Mann-Whitney
Paired t-test	Histogram of paired differences	Wilcoxon signed rank

ANOVA

Compares the means of several groups

Which diet is best?

Dependent: Weight lost (Scale)

Independent: Diet 1, 2 or 3 (Nominal)

Null hypothesis: The mean weight lost on diets 1, 2 and 3 is the same

$$H_0 : \mu_1 = \mu_2 = \mu_3$$

Alternative hypothesis: The mean weight lost on diets 1, 2 and 3 are not all the same

Post hoc tests

If there is a significant ANOVA result, pairwise comparisons are made

They are t-tests with adjustments to keep the type 1 error to a minimum

- Tukey's and Scheffe's tests are the most commonly used post hoc tests.
- Hochberg's GT2 is better where the sample sizes for the groups are very different.

Assumptions for ANOVA

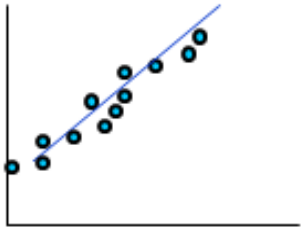
Assumption	How to check	What to do if assumption not met
Normality: The residuals (difference between observed and expected values) should be normally distributed	Histograms/ QQ plots/ normality tests of residuals	Do a Kruskall-Wallis test which is non-parametric (does not assume normality)
Homogeneity of variance (each group should have a similar standard deviation)	Levene's test	Welch test instead of ANOVA and Games-Howell for post hoc or Kruskall-Wallis

Correlation Coefficient r

- **Measures strength of a relationship between two continuous variables**

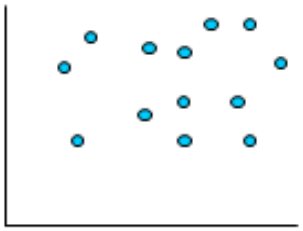
$$-1 \leq r \leq 1$$

Strong positive linear relationship



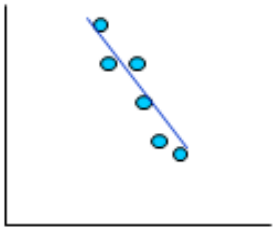
$r = 0.9$

No linear relationship



$r = 0.01$

Strong negative linear relationship



$r = -0.9$

Summary Table of Statistical Tests

Level of Measurement	Sample Characteristics					Correlation
	1 Sample	2 Sample		K Sample (i.e., >2)		
		Independent	Dependent	Independent	Dependent	
Categorical or Nominal	X^2 or bi-nomial	X^2	Macnarmar's X^2	X^2	Cochran's Q	
Rank or Ordinal		Mann Whitney U	Wilcoxin Matched Pairs Signed Ranks	Kruskal Wallis H	Friendman's ANOVA	Spearman's rho
Parametric (Interval & Ratio)	z test or t test	t test between groups	t test within groups	1 way ANOVA between groups	1 way ANOVA (within or repeated measure)	Pearson's r
		Factorial (2 way) ANOVA				

(Plonskey, 2001)