

WHAT STATISTICAL TEST DO I NEED?

Deciding on appropriate statistical methods for your research:

What is your research question?

Which variables will help you answer your research question and which is the dependent variable?

What type of variables are they?

Which statistical test is most appropriate? Should a parametric or non-parametric test be used?

KEY WORDS:

VARIABLE: Characteristic which varies between independent subjects.

CATEGORICAL VARIABLES: variables such as gender with limited values. They can be further categorised into **NOMINAL** (naming variables where one category is no better than another e.g. hair colour) and **ORDINAL**, (where there is some order to the categories e.g. 1st, 2nd, 3rd etc).

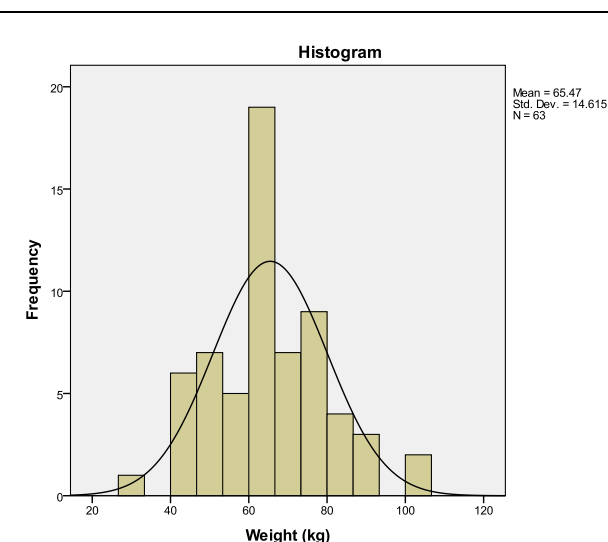
CONTINUOUS (SCALE) VARIABLES: Measurements on a proper scale such as age, height etc.

INDEPENDENT VARIABLE: The variable we think has an effect on the dependent variable.

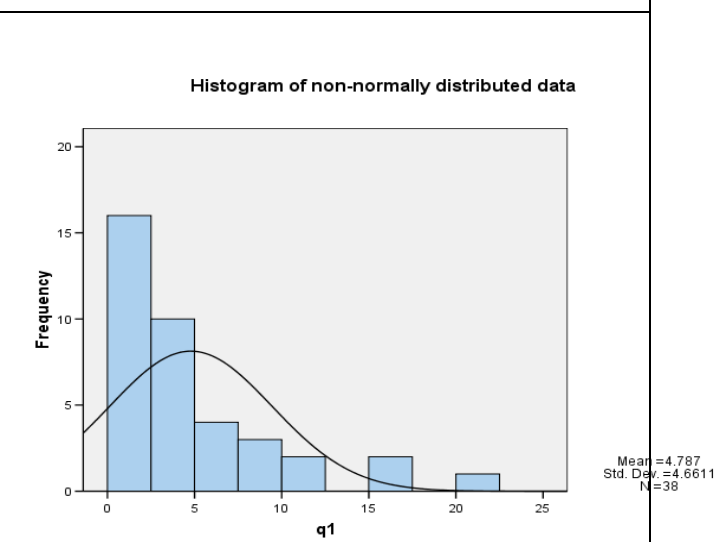
DEPENDENT VARIABLE: The variable of interest which could be influenced by independent variables.

PARAMETRIC TESTS: there are various assumptions for parametric tests including the assumption that continuous dependent variables are normally distributed. There are specific tests for this within packages such as SPSS but plotting a histogram is also a good guide. As long as the histogram of the dependent variable peaks in the middle and is roughly symmetrical about the mean, we can assume the data is normally distributed (see examples below).

Example of data which is approximately normally distributed



Example of skewed data



Comparing:	Dependent variable	Independent variable	Parametric test (Dependent variable is normally distributed)	Non-parametric test
The means of two INDEPENDENT groups	Continuous/scale	Categorical/nominal	Independent t-test	Mann-Whitney test
The means of 2 paired (matched) samples e.g. weight before and after a diet for one group of subjects	Continuous/scale	Time variable (time 1 = before, time 2 = after)	Paired t-test	Wilcoxon signed rank test
The means of 3+ independent groups	Continuous/scale	Categorical/nominal	One-way ANOVA	Kruskal-Wallis test
The 3+ measurements on the same subject	Continuous/scale	Time variable	Repeated measures ANOVA	Friedman test
Relationship between 2 continuous variables	Continuous/scale	Continuous/scale	Pearson's Correlation Co-efficient	Spearman's Correlation Co-efficient (also use for ordinal data)
Predicting the value of one variable from the value of a predictor variable	Continuous/scale	Any	Simple Linear Regression	
Assessing the relationship between two categorical variables	Categorical/nominal	Categorical/nominal		Chi-squared test

Note: The table only shows the most common tests for simple analysis of data.

Examples:

- **Are height and weight related?** Both are continuous variables so Pearson's Correlation Co-efficient would be appropriate if the variables are both normally distributed.
- **Can height predict weight?** You cannot determine height from weight but you could estimate weight given height so height is the continuous independent variable. Simple linear regression will help decide if weight is a good predictor of height and produce an equation to predict weight given an individual's height.
- **Is Diet 1 better than Diet 2?** A researcher would randomly allocate subjects to two groups with one group following Diet 1 and the other Diet 2. Weight would be taken before and after the diet and the mean weight lost compared for the two groups. The dependent variable 'weight lost' is continuous and the independent variable is the group the subject is in which is categorical. If the data is normally distributed, use the independent t-test, if not use the Mann-Whitney test.
- **Are patients taking treatment A more likely to recover than those on treatment B?** Both 'Treatment' (A or B) and 'Recovery' (Yes or No) are categorical variables so the Chi-squared test is appropriate.