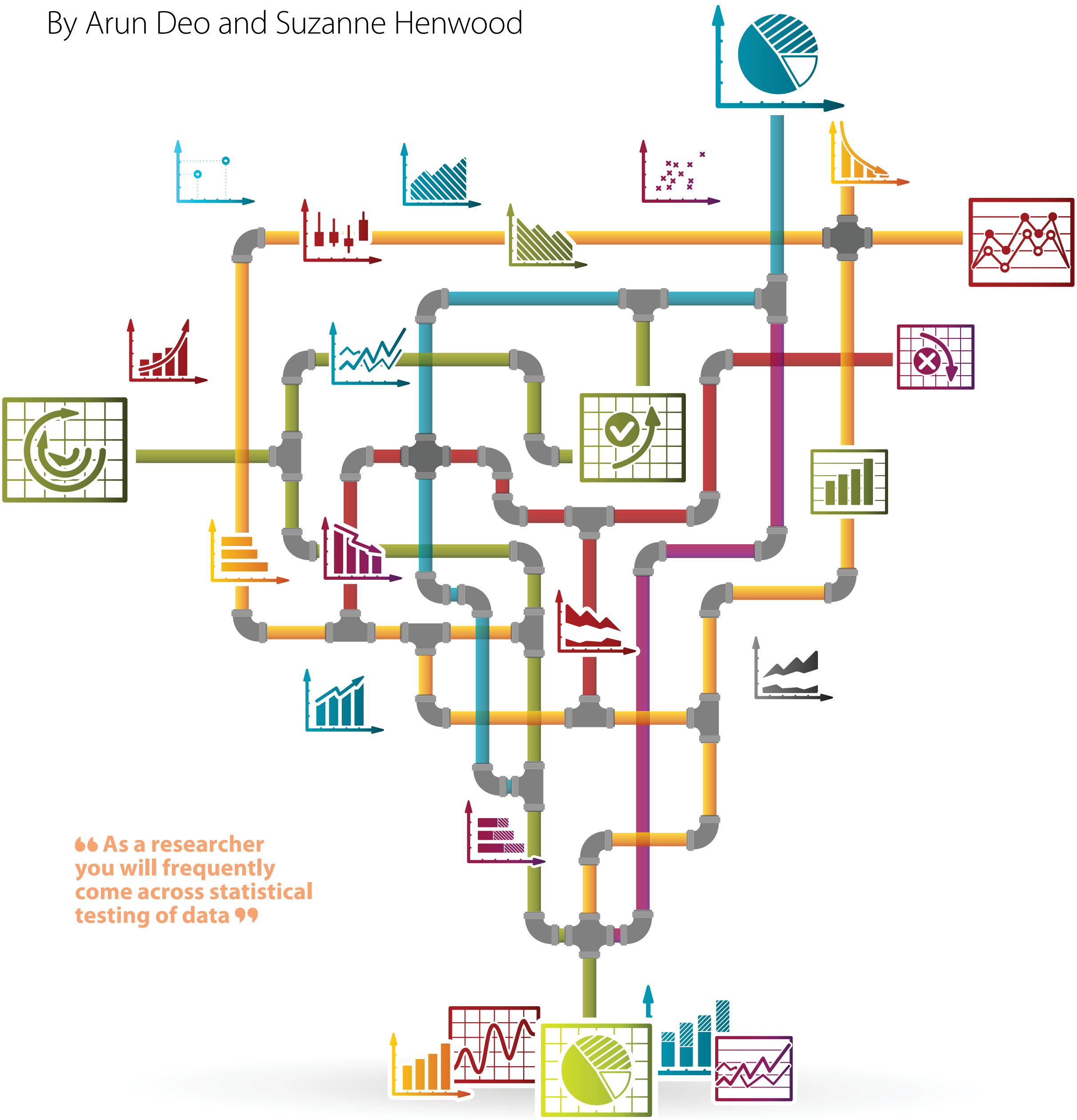
Introduction to Statistics in Research

By Arun Deo and Suzanne Henwood



“As a researcher you will frequently come across statistical testing of data”

As a researcher (or consumer of research) you will frequently come across statistical testing of data. This article aims to give an introductory outline to statistical methods to enable you to (i) choose appropriate statistics for your own study and (ii) offer critical comment on statistics used in journal articles or research papers.

This background knowledge will give you the ability to comment on research design adequacy, which can lead to you being able to know whether you can trust and use results in practice, or guide you in critical questioning of the validity and reliability of results, leading you to search for further evidence before implementing new suggestions into your practice.

This article is then aimed at exploring the use of statistics in quantitative studies, generally from data collected either through (i) survey data collection or (ii) criteria measurements. Criteria measurements in NLP might for example include measurement of anxiety or depression, pre and post an NLP intervention (ideally using a pre validated measurement tool which has been shown to be valid and reliable – i.e. has been shown to be accurate and repeatable over time). Another example would be looking at spelling test scores following work with pupils on the NLP Spelling Technique.

Key terminology

Some key terms which would be helpful to define up front are as follows.

Population: this is the whole group of people from which you take a *Sample* to make research more manageable.

Let's put this into perspective using the NLP spelling example.

The population would be all pupils in the school involved in the study.

The sample could be a random selection of those pupils, who would be offered the NLP spelling strategy as a workshop. (The precise number in the sample is ideally determined by a power calculation to make the inferential statistics valid. Tools to assist with this are available on the web.) Other sampling options are also available including involving the total population: or selecting through, for example, stratification across a variable you think might impact on results (e.g. age or previous test results), or through asking for volunteers (though this may be biased as those who respond may be part of a specific group of children, e.g. very keen or very poor spellers). The sampling used is useful to look at when assessing if results are valid for a total population.

Variables: dependent and independent

Variables can take on more than one value (*1, p. 9) and in this case might include the age of the student, their gender, previous grade average.

The independent variable would be the way the spelling was taught in the study (i.e. those offered NLP Spelling strategy versus those taught only by traditional means and maybe even some children who were not specifically taught how to spell who are included as a control group). The dependent variable would be the test scores.

Types of data

Nominal data – has no attached meaning. Its value does not signify anything (*1, p. 10), it is just a label. This might, for example, be the room the student is in (room 5).

Ordinal data – is where numbers start to take on meaning. For example, a fear of 7 out of 10 is greater than a fear of 3 out of 10. However, it does not mean the difference between 3 and 4 is the same as the difference between 7 and 8.

Interval data – represents data where the interval between numbers is equal (and where zero means the complete absence of whatever you are measuring). So a score of 3 correct spellings out of 10 is half a score of 6 correct spellings out of 10 and zero would mean none were correct.

“Non-parametric methods are sometimes referred to as distribution-free”

Ratio data – includes a meaningful zero point (*1, p. 11), for example a measurement using a ruler.

Parametric statistical methods – make an assumption that there is an underlying normal distribution of what is being measured (i.e. across the parameters of variables being measured). (It assumes then that the occurrences, if measured across the whole population, would show a normal distribution curve – like a bell shaped curve on a graph.)

Non-parametric statistical methods– can be used on data with any distribution.

Note: where it is known that a normal distribution exists, greater inference and rigour can be applied by using parametric tests, allowing comment to be made on the whole population (but if they are used on non-parametric data, measurements can be misleading). Non-parametric tests also make it more difficult to allow for confounding factors using multiple regression (which might be applied in more advanced statistics).

Parametric methods

Let's start looking at how we use these statistical methods in practice. Firstly, let's look at parametric methods of statistical tests.

Figure 1 is a flow chart to outline a wide range of possible tests, across numerous types of data, which we will use to explore statistical options. The branches under 'Yes' correspond to tests associated with true quantitative data (that is, ordinal, interval and ratio scales: numbers which are meaningful, that measure something on a scale, e.g. anxiety, weight, blood pressure etc). The 'No' options correspond to tests

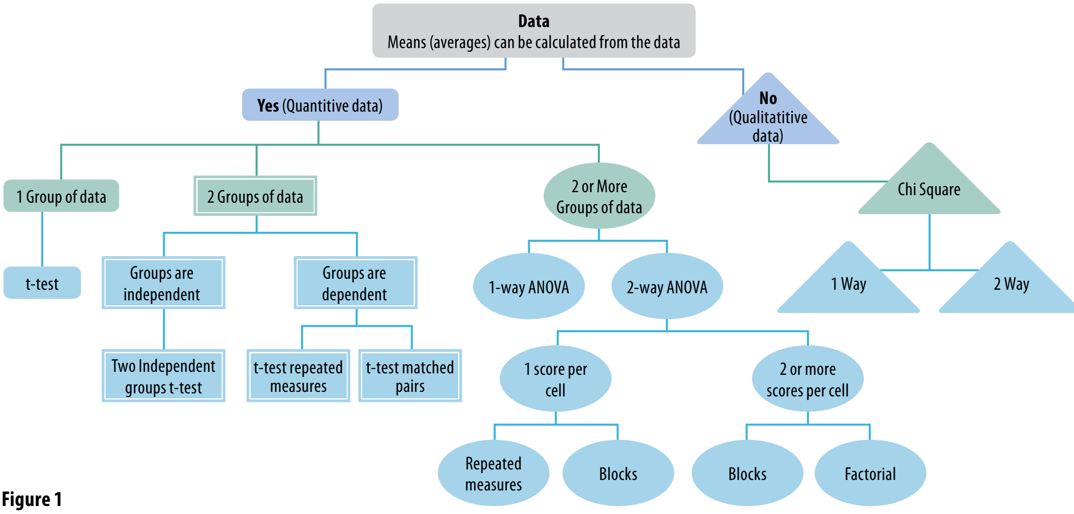


Figure 1

associated with qualitative data which is either/or (e.g. gender, smoking status) or just data labels (e.g. class number), often called nominal or dichotomous data. (Note: this is different from qualitative methodologies, which focus on text based or interview data.)

Working through the flow chart: a t-test is a statistical hypothesis test in which the test statistic follows a student's t distribution if the null hypothesis is supported. (The null hypothesis is that there is no significant difference demonstrated.)

Two independent groups' t-test is an inferential statistical test that determines whether there is a statistically significant difference between the means in two unrelated groups.

The t-test for repeated measures and matched pairs called the two dependent group t-test (also called the paired t-test) compares the means of two related groups to detect whether there are any statistically significant differences between these means.

The *one-way* analysis of variance (ANOVA) is used to determine whether there are any significant differences between the means of three or more independent (unrelated) groups.

The *two-way* analysis of variance (ANOVA) is an extension of the one-way ANOVA that examines the influence of two different categorical independent variables on one continuous dependent variable.

The repeated measures/blocks/factorial ANOVA are more advanced statistical tests associated with two-way ANOVA.

A chi-square test is a statistical hypothesis test associated with categorical variables to measure the goodness of fit or test for independence or to find out if the sub groups are homogeneous.

Non-parametric methods

Non-parametric methods have been developed for studies in which the assumptions necessary for using parametric methods cannot be made. Non-parametric methods are sometimes referred to as distribution-free methods because it is not necessary to assume that the observations are normally distributed. A non-parametric method is appropriate for dealing with data that are measured on a nominal or ordinal scale and whose distribution is unknown. Because of the many advantages of non-parametric methods, their use has been increasing rapidly. But like most methods, they also have disadvantages.

> **“A starting point to designing your own studies, testing data and reading research critically”**

Non-parametric methods have three main advantages (*2, p. 259):

1. They do not have restrictive assumptions such as normality of the observations. In practice, data are often non-normal or the sample size is not large enough to gain the benefit of the central limit theorem. At most, the distribution should be somewhat symmetrical. This gives non-parametric methods a major advantage.
2. Computations can be performed speedily and easily – this gives a prime advantage when quick preliminary indication of results is needed.
3. They are well suited to experiments or surveys that yield outcomes that are difficult to quantify. In such cases, the parametric methods, although statistically more powerful, may yield less reliable results than the non-parametric methods, which tend to be less sensitive to errors inherent in ordinal measurements.

There are also three distinct disadvantages of non-parametric methods (*2, p. 259):

1. They are less efficient than comparable parametric tests (i.e. they require a larger sample size to reject a false hypothesis).
2. Hypotheses tested with non-parametric methods

are less specific than those tested comparably with parametric methods.
3 They do not take advantage of all of the special characteristics of a distribution. Consequently, these methods do not fully utilise the information known about a distribution.

In using non-parametric methods, you should be careful to view them as complementary statistical methods rather than attractive alternatives. With a knowledge of their advantages and disadvantages, and some experience, you should be able to determine easily which statistical test is most appropriate for a given application (i.e. try to use parametric if at all possible). This understanding will also enable you to ask relevant questions about choice of statistics in others' work.

There are numerous non-parametric methods. In this introductory article we have limited ourselves to those that correspond to parametric t tests for single samples, independent samples, dependent or paired samples, and correlation coefficients. We also present Fisher's exact test, which is appropriate when the chi-square test would not be valid to use. These techniques are the Wilcoxon rank-sum test/Mann-Whitney U test, the Spearman rank-order correlation coefficient, the Kruskal-Wallis one-way ANOVA, and the sign test. Figure 2 lists parametric tests and their non-parametric equivalents.

Knowing the options available should make it easy to go and search for further information on any method you want to explore.

Summary

This brief overview of statistical methods is offered to give a first look at quantitative data management. It cannot cover all methods in full, but is offered to highlight what options are available for basic data sets and signposts where to search further for particular types of data by giving the names of statistical methods to look up on the web or in appropriate text books. We hope it is a helpful starting point to designing your own studies, testing data and reading research critically as you build your confidence in handling statistics and using research in practice.

	Parametric test	Non-parametric test
One sample	One-sample t test	• One-sample sign test
Two independent samples	Two-sample independent t test	• Wilcoxon rank-sum test • Mann-Whitney U test
Two dependent samples	Two-paired t test	• Wilcoxon signed-rank test • Sign test
Correlation	Pearson r	• Spearman rank-order correlation
Multiple groups, one factor	One-way ANOVA	• Kruskal-Wallis one-way ANOVA

Figure 2

References
(*1) J Schmuller, *Statistical Analysis with Excel for Dummies*.
(*2) JW Kuzma and SE Bohnenblust, *Basic Statistics for the Health Sciences* (5th edition).

Arun Deo: Biostatistician, Unitec Institute of Technology, Auckland.
Suzanne Henwood: Associate Professor in Social and Health Sciences, Unitec Institute of Technology, Auckland.