

Inferential statistics - 2

Buddhika K Dassanayake

Department of Surgery
Faculty of Medicine, University of Peradeniya

Key words:

Introduction

This final section of the series will deal with the many tests by which groups may be compared with each other. Since many of the statistical methods employed will be familiar by name to the readers, this article will focus more on the appropriate application of each method. Any menu-driven software (such as SPSS or MINITAB) will allow a user to carry out the tests given below. Therefore we will deal with selecting the appropriate tests.

In keeping with the structure of the series, we will deal with the statistical methods available to us depending on the type of data that we are analyzing.

1. Categorical data

Sometimes, we have to analyze associations between variables that are categorical.

For example, a hospital may conduct a survey as to how many of its in-patients are smokers and how many are non-smokers, and also as to how many patients have cancer and do not have cancer. The task is upto us to see if there is an association between smoking and cancer. Out of 154 cancer patients, 124 are nonsmokers. And out of 121 patients without cancer, 89 are smokers. How do we go about this?

The first step in analyzing categorical data is to create a contingency table

Eg:

This gives us a good summary of the data.

Column1	Smoker	Nonsmoker
Cancer	124	30
Healthy	32	89

Correspondence : Dassanayake BK
Email : thraless@gmail.com

The second step is to select an appropriate statistical test for comparing categorical data

The two commonly used tests are **Fisher's Exact Test** and **Chi Squared Test**

Fisher's exact test is limited to examples such as that above, where each variable has only two levels (ie, it cannot be used if the smoking status had three levels such as 'smoker', 'ex-smoker' and 'nonsmoker'.) However it gives more valid results when the count in each cell is a small number.

Chi Squared test can be used even when one variable has more than two levels. However, it gives a less accurate result when the count in a cell is a small number.

Both tests give a p value as their result. A small p value means that there is a significant association between the variables.

Two other statistical tests that are useful in analyzing categorical data are the **odds ratio** and **relative risk**. Similar to the Chi Squared or Fisher's tests, these statistics will also give a p-value which indicates whether the association is significant or not. However, they have the added advantage of **quantifying the risk or odds of one group having the condition in comparison with the other**. The Odds Ratio is commonly used in Case-control studies, while Relative Risk is commonly used in Cohort Studies.

2. Ratio / interval data

We are commonly required to do four types of analysis using ratio/interval data:

- compare a variable between two different groups
- compare a variable in the same group before and after a treatment
- compare a variable between more than two groups

d. compare two different variables with each other

a. comparing a variable between two different groups

Eg: comparing the birth weights of Sri Lankan and British neonates

The simple test of choice would be an **independent sample t test (also called a pooled t-test)**

b. comparing the same group before and after a treatment

Eg. Weight of a group of children before and after Viral fever.

Time taken for subjects to complete a task before and after administering sedative drugs.

The simple test of choice would be a **paired t-test**

c. comparing a variable between more than two groups

Eg. Difference in wound-healing time between four types of wound dressings.

The simple test of choice would be an **analysis of variance (ANOVA)**.

The principle behind an ANOVA is a simple one. The question asked is “*Is the difference between these groups due to the treatment, or simply due to random chance?*”. In statistical terms, this may be stated as a “comparison of treatment effect with the effect of random error”. The null hypothesis thus assumes that any difference in wound healing time between the four types of dressing is due to random error. Therefore, **if the p value of an ANOVA is very low (<0.05), we can reject the null hypothesis and say that different treatments significantly alter wound healing time.**

However, a significant result only means that *at least two of the four treatments are different from each other*. Further testing (beyond the scope of this overview) will have to be done to determine which ones actually are (Mean Separation Techniques).

d. comparing two different variables with each other

There are two common methods by which this is done: using **Pearson's coefficient of correlation** and by **regression**.

Pearson's coefficient (r) is a value which may vary from -1 to +1. A negative value shows a negative

correlation (ie: one variable increases as the other decreases). The closer the value is to 0, the weaker the correlation. It also comes with a p value. Thus a good correlation should be both strong (closer to 1) as well as significant.

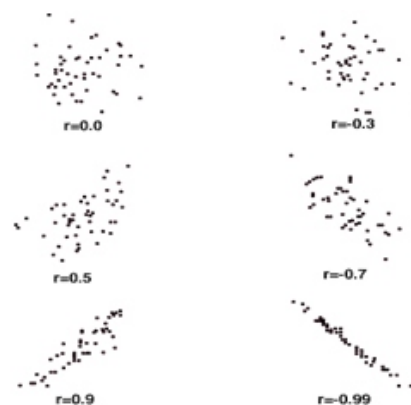


Fig. 1: various associations and their coefficients of correlation

Regression is another useful method which not only lets us compare two variables and see how they are related, *but also allows us to predict how one variable will affect the other.*

To understand regression, first let us study the simple linear equation $y = mx$. The graph would be a perfect straight line as shown in Fig. 2:

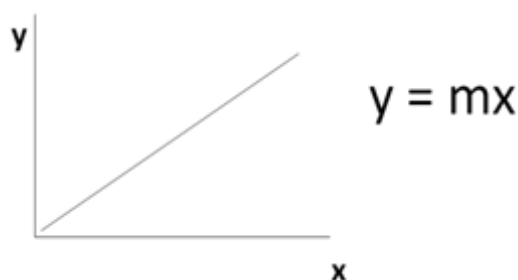


Fig. 2: $y = mx$

However, no relationships in real life are that perfect! In reality, any such graph drawn between two variables (Eg: the height and weight of adult men) would be closer to that in Figure 3:

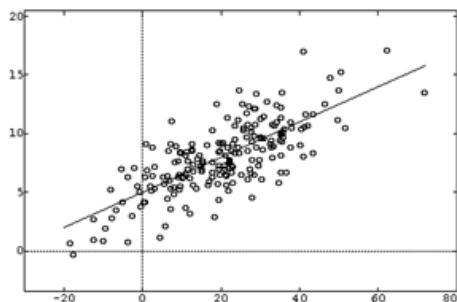


Fig. 3: a scatter plot between two variables with an equation line drawn through it

As you can see, all the points do not fit on the line. Yet we can draw an equation line through it somewhat like $y = mx$. Some of the variability can be explained by our equation. The rest of the variability is due to random error.

Thus, for example, if we were comparing heights and weights of men, Regression attempts to calculate **how much of the change in weight is due to the change in weight? And how much in the change in weight is due to random error?**

Regression calculations give us a value known as the **R² value. R² is the proportion of the variability of y that can be explained by x.** Therefore, the closer R² is to 1, the more significant the association.

Thus our formula becomes something like $y = mx + E$ where E is the random error.

For example, let us assume that we wish to find the factors which affect Peak Expiratory Flow Rate. We will first collect data including height, age, etc of subjects and then measure their peak flow rates. Next, we will create 'Regression Equations' such as this:

$$\text{Peak flow rate} = a \cdot \text{height} + b \cdot \text{age}$$

We then calculate the R² value, and keep modifying the equation until we reach the maximum possible R² value, which may be with an equation such as this:

$$\text{Peak flow rate} = a \cdot \text{height}^2 + b \cdot \text{age}^3$$

Thus, we now have an accurate **predictive equation**, where we can calculate the expected peak flow rate for

any person if we know their age and height.

3. Ordinal data

We sometimes have to carry out the same kinds of analysis with Ordinal data. Fortunately, **each type of test for ratio/interval data has an analogous ordinal test.**

These are known as **non-parametric tests.** In addition to being useful in analyzing ordinal data non-parametric tests *are also safe to use even with ratio data with small sample sizes.*

The table below compares the necessary tests for each type of data.

To summarize the content of this series, any study may be broken into a series of ordered steps:

- What do I want to know (ie the research question)?
- What type of data will I be using?
- What is my sample size?
- What are my descriptive statistics?
- What tests will I perform on the data to answer my question?

References (also to be used as further reading)

- Hill, AB. Principles of medical statistics. 9th ed. London: The Lancet; 1971.
- Cochrane AL. Effectiveness and Efficiency. Random Reflections on Health Services. London: Nuffield Provincial Hospitals Trust, 1972.
- Thattil RO, Gunaratne LHP. Basic Statistics for Advanced Level Students. Colombo: Godage & Bros; 2000.
- Dodge, Y. (Ed.). The Oxford Dictionary of Statistical Terms. Oxford: The Oxford University Press; 2003.
- Stevens, S. S. "On the Theory of Scales of Measurement." Science, New Series, Vol. 103, No. 2684 (Jul. 7, 1946), pp. 677-680

<i>What do I want to do</i>	<i>Ratio/Interval data</i>	<i>Ordinal data</i>
Compare two different groups	Independent sample t test	Mann-Whitney u test
Compare a group before and after a treatment	Paired t test	Wilcoxon signed rank test
Compare two different variables	Pearson's coefficient	Pearson's coefficient
Compare more than two groups	ANOVAs	Kruskal-Wallis or Friedmann tests