**Why do we use statistics?**

In simple terms we use statistics in order to eliminate grey areas in our data. It is sometimes very clear whether a data set is correlated, or whether there is an association between two variables (i.e. House prices and location), however when you statistically test a data set you can reach a certainty level which allows you the eliminate the fact that your data set has occurred by chance.

If you think of it as an analogy, we use statistics like a drunken man uses a lamppost – for support and not for illumination.

In other words doing a statistical test isn't the be all and end all of life, but doing a stat's test will help you support or prove your conclusions.

All you need to know now is which one to use, and when…!!

**Spearman's Ranks and Pearson's Product:**

Spearman's Ranks Tests and Pearson's Product Co-efficient statistically tests the strength of a correlation, i.e. as one variable increases/decreases does the other increases/decreases.

**Chi$^2$:**

A Chi$^2$ Test tests association. It looks at whether there is a difference in the values you have collected and the values you would mathematically expect. You use a Chi if your data is split into categories, i.e. age ranges (0-10, 11-20), sizes (small, medium, large), or feeding levels (herbivore, detritivore, carnivore etc)

**T-Test:**

T-tests are used to establish whether there are DIFFERENCES in data sets. I.e. there is a difference in species diversity in Light and Dark areas, or Grazed and Non-Grazed areas etc. A T-Test can look quite complicated however it is a good statistical test to use as it accounts for variance (the spread of numbers) within a data set.

**Mann-Whitney U:**

Mann-Whitney U is also used to establish whether there is a DIFFERENCE between data sets, however it does not take into account variance, unlike a T-Test.

**The Hypotheses**

All statistical tests start with a **Null** hypothesis, and an **Alternative** hypothesis. The reason for this is much like the UK criminal justice system; you are innocent until proven guilty. In stats you have to accept there is no statistical relationship between a data set, unless you can prove otherwise. That's why you start with a **Null hypothesis** which is state: There is NO statistical Correltion/Association/Difference between…. You are always trying to disprove your Null hypothesis. If you beat any of the tests above you can do this by Rejecting the Null hypothesis and Accepting the Alternative hypothesis.

**Spearman's rank**

Ok…so a spearman's rank is simple. You conduct a spearman's rank test when you are trying to establish the strength of a correlation between two variables. This may sound complicated but it really isn't. Variables are just things that you measure. A variable could be temperature, gradient, pH, Cross-sectional Area, soil moisture, altitude, or any other number of things.

The easiest way to establish whether you can do a spearman's rank test or not is to plot your data on a scatter graph, if you can plot your data on a scatter graph you can more than likely conduct a spearman's rank on it.
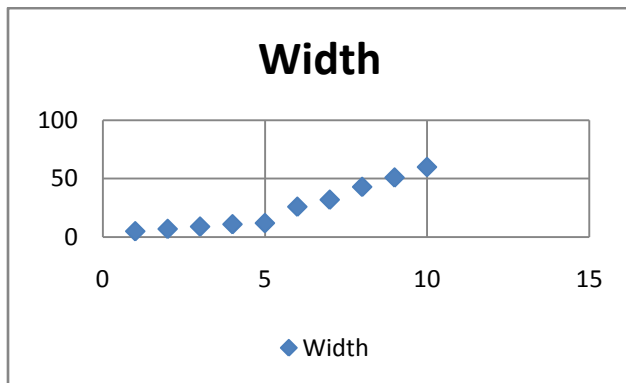
**Graph 1**                                    **Graph 2**
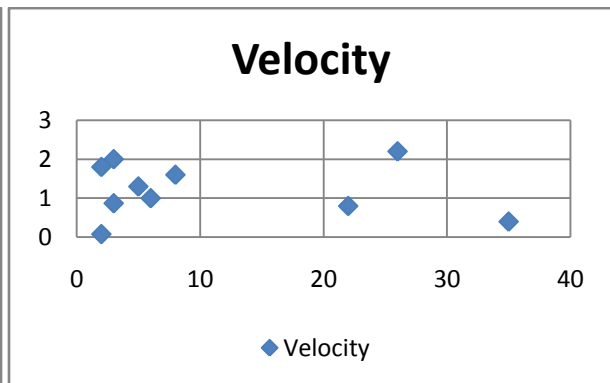


Above are two scatter graphs. The first is showing the relationship between distance downstream and width in a river. It is easy to see by the graph that the width is increasing the further we go downstream, however the second graph is less clear. Graph 2 is showing the relationship between Velocity and Pebble size. Although it may look like there is no correlation at the moment, by doing a spearman's rank on the data set it is possible to establish whether there is a positive or a negative correlation, and also how strong the correlation is (See 'How to conduct a Spearman's rank test under the 'Statistical Resources' Tab to conduct your own).

One disadvantage of a Spearman's Rank Test is that it does not take into account variance within the data set. The test works on whether the ranks of each data set relate to each other. Once you have ranked the data set the test doesn't care what the original numbers were, it just works from the ranks you have given the data.

**Pearson's Product**

A Pearson's Product test can also be used to test the strength of a correlation between two variables. The number you get out at the end of the test will also fall between +1 and -1 telling you whether the correlation is positive or negative. A Pearson's however does take into account variance within the data sets. Instead of ranking the data sets the test works from the actual numbers you have collected instead of the ranks which you have given them.

**Chi²**

A Chi² test looks at association between variables. The test is also used when the data which has been collected is in categories, at the end of the test the number obtained is compared to a critical value table and if the number is higher than the critical value a degree of certainty can be reached as to whether there is a statistical association between the variables.

The test works relatively simply but comparing the values you have collected to those you would mathematically expect to see. I.e. if there were 10 balls in a box, some were red, and some were blue, mathematically you would expect there to be 5 of each colour. To start the test you enter the 'Observed' values, which are the values you actually obtained. Then you work out the 'Expected' values, these are what you would mathematically expect to see based on the numbers in the data set.

To work out the expected values the equation is **Row Total x Column Total / Grand Total.**

Categories    Expected

Observed    Observed - Expected²

Observed    Observed - Expected

| Feeding level | | River | | | | | Pond | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | O | E | O-E | $(O-E)^2$ | Chi Value | O | E | O-E | $(O-E)^2$ | Chi Value | Row Total |
| Herbivores | 58 | 51.35 | 6.65 | 44.21 | 0.86 | 23 | 29.65 | -6.65 | 44.21 | 1.49 | 81 |
| Detritivores | 35 | 57.69 | -22.69 | 514.86 | 8.92 | 56 | 33.31 | 22.69 | 514.86 | 15.46 | 91 |
| Carnivores | 75 | 58.96 | 16.04 | 257.33 | 4.36 | 18 | 34.04 | -16.04 | 257.33 | 7.56 | 93 |
| Column Total | 168 | | | | 14.15 | 97 | | | | 24.51 | 265 |
| | | | | | | | | | | Chi Value | 38.66 |

In the example above you can see that there were 58 Herbivores observed in the river, however you would mathematically expect there to have been 51.35. The equation used for the 'Expected' values basically looks at how many herbivores you had at both sites (River and Pond), multiplies this number by how many organisms you had in total at the river, and then divides it by the number of organisms you had in total at both sites. This then tells you how many Herbivores you would mathematically expect there to be based on the data set collected. You then repeat this for all of the other categories. The column after this then looks at the difference between the Observed and Expected values. The bigger the difference between the two, the more likely there is to be an association between the variables. The best example from the table above is the detritivore column, 35 detritivores were observed at the River, however you would mathematically expect there to be 57.69. This is quite a big difference. Therefore there is probably going to be an association between feeding level and habitat.

**T-Test**

A T-Test looks at whether two data sets are different. The test also looks at the Standard Deviation and Variance of the data set, i.e. how much the numbers fluctuate around the mean. The final t-test equation is basically the difference of the means of both data sets divided by the variance of both sets added together. This means the t-test can establish whether there is a substantial difference between the average values, whilst also accounting for variations in the data sets themselves. If you the example of Species Diversity scores, you may want to test whether the species diversity is different in Light and Dark areas of woodland or something similar to this. The t test will see whether the average species diversity for each site is representative of the data set as a whole, and also tell you whether the species diversities in both areas are different enough to be statistically significant.

The fact that it takes into account variance is very important as the mean value of a data set may not always be representative of the data set as a whole. For example the mean value of 100, 3, and 1 is 34.6, the standard deviation of these numbers is 56.59, and the variance is 3202.4. By looking at the high variance and standard deviation it is possible to see just how unrepresentative the mean is of the whole data set. T-Tests take into account the spread of data which is what makes them such good statistical tools. You can also use a t-test with unequal data sets. This means if you had a range of 15 species diversities for 1 area, and only 10 for another you could still use the test.

**Mann-Whitney U test**

A Mann-Whitney U test also looks at differences between data sets, however unlike a t-test it does not take into account variance within the data sets. It works from a ranking system which basically tests whether the ranks of two data sets are different or not. You rank the whole data set as a whole, i.e. let's say you are studying the height of plants in light and dark areas, if you had 20 samples, 10 from light areas, and 10 from dark areas; you would rank all 20 samples as a whole. It then looks at whether one of the areas got mainly low ranks, and the other got given high ranks. The example below shows a perfect difference between two data sets. One got all the low ranks, and one got all the high ranks, therefore the numbers are different.

Low Ranks
1 2 3 4 5 6 7 8 9 10

Difference

High Ranks
11 12 13 14 15 16 17 18 19 20
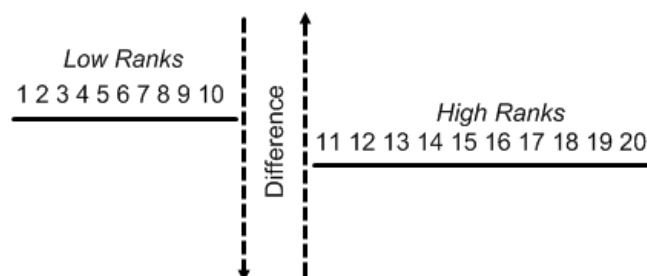
As mentioned before, statistics are mainly used to eliminate grey areas in our data. The example above would work like so:

Light area plant heights (cm): 10 11 12 13 14 15 16 17 18 19

Dark area plant heights (cm): 30 31 32 33 34 35 36 37 38 39

If you were ranking this data set as a whole you would end up with this:

*Ranks*   1  2  3  4  5  6  7  8  9  10
Light area plant heights (cm): 10 11 12 13 14 15 16 17 18 19

Dark area plant heights (cm): 30 31 32 33 34 35 36 37 38 39
*Ranks*  11 12 13 14 15 16 17 18 19 20

*Low Ranks*
1 2 3 4 5 6 7 8 9 10

Difference

*High Ranks*
11 12 13 14 15 16 17 18 19 20
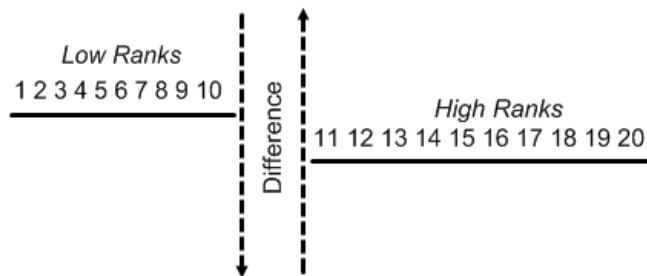
In real life a data set would probably over lap a little bit. This is when a Mann-Whitney U is useful as it would tell you whether the overlap prevents the data sets from being significantly different.