

ELEMENTARY STATISTICS

Statistics are useful mathematical tools which are used to analyse data. Perhaps the best known statistic is the average. This is a single figure which is used to represent a set of data.

Averages

There are three types of average:

The median which is the middle value of a range of results.

The mode which is the value that appears the greatest number of times.

The mean which is the sum of all the results divided by the number of results.

Example: in the following set of data: 1; 3; 7; 10; 11; 12; 13; 13; 22; 23; 24

The median = 12 which is the 6th value when they are arranged in order.

The mode = 13

The mean = $139 \div 11 = 12.6$

Note that if there are an even number of results the median is calculated by adding the middle two values and dividing by two.

Example: for the series 1; 2; 4; 5; 7; 9 the median = $(4+5)/2 = 4.5$

If all three averages are approximately the same we can usually assume that the data shows a “**normal distribution**” and the following tests can be used even for comparatively small samples ($n < 30$).

Science usually works by taking samples rather than measuring every last item in a population.

The mean of the sample you have selected is not necessarily the mean of the whole population. Nor is it necessarily true that if the sample means taken from two different populations are different then the population means of each must be different. There is bound to be a natural variation. The tests below determine how big the variation you obtain is in a sample.

Measuring the spread of the data

Averages do not tell us everything about a sample. Samples can be very uniform with the data all bunched around the mean or they can be spread out a long way from the mean. The statistic that measures this spread is called the standard deviation.

Arrange your data as follows and carry out the calculations:

Observations e.g.Number of plants per quadrat	Frequency e.g.Number of quadrats		
x	f	xf	x²f
0			
1			
2			
3			
4			
5			
etc.			
	Σf = n	Σxf	Σx²f

(Note: Σ = sum of)

The mean of the sample: $\bar{x} = \frac{\sum xf}{n}$

The standard deviation of the sample: $s = \sqrt{\frac{\sum x^2f - \frac{(\sum xf)^2}{n}}{n - 1}}$

You can use a programmable calculator or a spread sheet to do this calculation for you after you have entered the data.

The standard deviation is a measure of the variation of the results. For data that is evenly distributed each side of the mean (a normal distribution) 68% of the data lies within one standard deviation of the mean.