New A-level Biology

Statistics booklet

CONTENTS

About Statistics -

There are two main kinds of statistics.

Measuring ‘Spread’
(Descriptive statistics)

In the Biology A-level, we will use descriptive statistics to measure the ‘spread’ of data. Descriptive statistics are used to describe data. For example, if you were investigating the number of visitors to a beach in August (nice job if you can get it!), you might draw a graph to see how the number of visitors varied each day, work out the average number of visitors each day (using mean, mode or median), work out the range of visitor numbers each day. This would all be descriptive statistics. Descriptive statistics also involve using:

- **Range**
- **± 2 Standard Deviations from the mean**
- **± 1.96 Standard Errors of the mean**

Can **NOT** use the term “significant”

Using what we know to make inferences about what we don’t know (Inferential statistics)

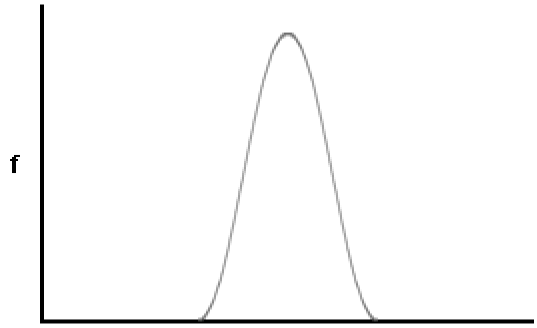
In the Biology A-level, we will use Inferential statistics which are techniques that allow us to use what we know to make inferences (i.e. judgements) about what we don’t know. For example, if we asked 200 people who they were going to vote for, on the day before the local election, we could try to predict which party would win the election. We need to choose a statistical test to use:

- **Chi squared**
- **Spearman’s rank correlation coefficient**
- **Student’s t-test**

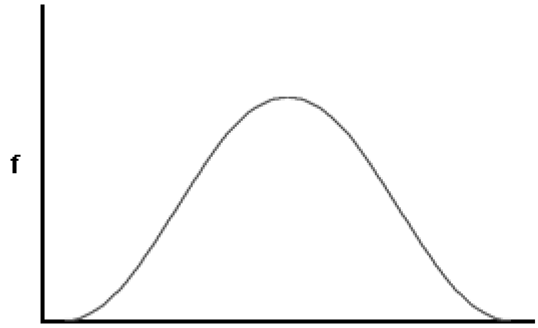
CAN use the term “significant”

<u>A note about using calculators</u>
We expect that you will often use electronic devices to calculate test statistics during your classwork.
In written examinations, you will **not** be asked to perform a calculation using a statistical test. It will be important for you to understand **how to select a statistical test** that is appropriate for given data and to be **able interpret the results** of such a statistical test. You could also be asked to explain your choices and interpretation.

Measuring ‘Spread’



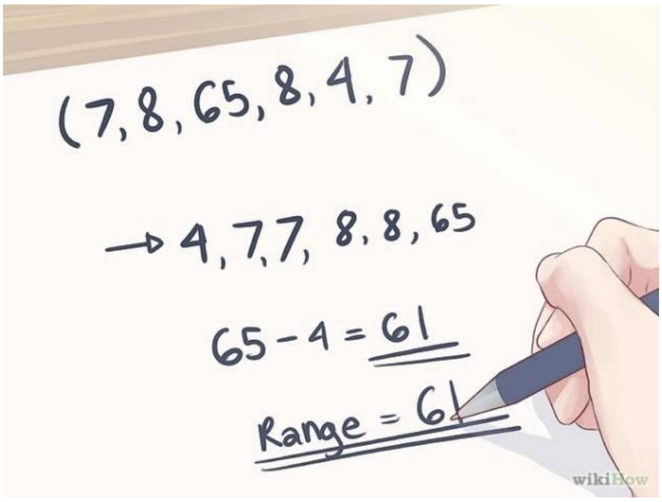
Distribution 1



Distribution 2

‘Distribution 1’ is tall and thin and ‘distribution 2’ is short and fat yet they have the same mean average (and this data also has the same ‘mode’ and same ‘median’ average too.) Nevertheless, the distributions have different spreads.

Range

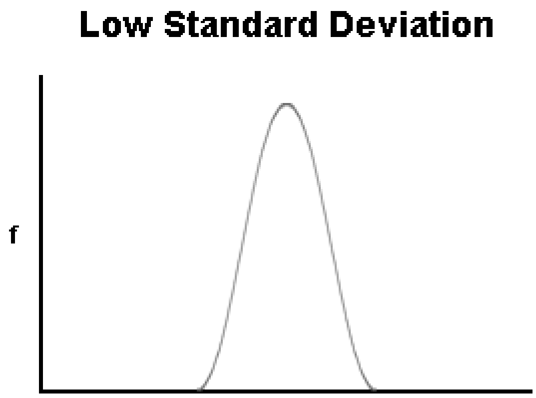


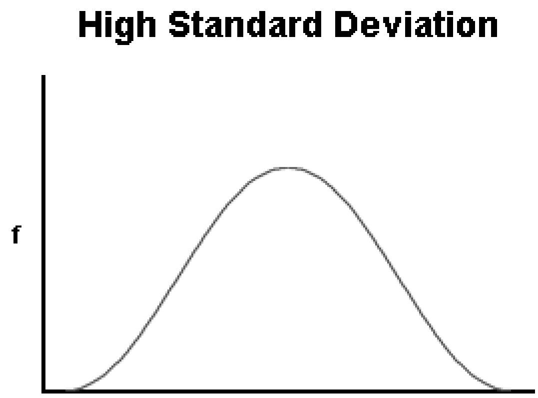
Range = largest value – smallest value

To measure the spread, we could calculate the range, but sometimes there are data which are **outliers.** Calculating range is simple but **often not a good measure of the spread**. Using the standard deviation as a measure of spread about the mean is often a better measure of spread.

Standard deviation

Using standard deviation is better than the range as it **uses all the observations**, and is less affected by the **outliers**.





A “thin” curve means that most values remain close to the average, and the standard deviation is small.

A “fat” curve means that there is a wider spread of values about the mean, and the standard deviation is large.

How to calculate Standard deviation

The standard deviation is calculated using the formula:

$$SD = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n-1}}$$

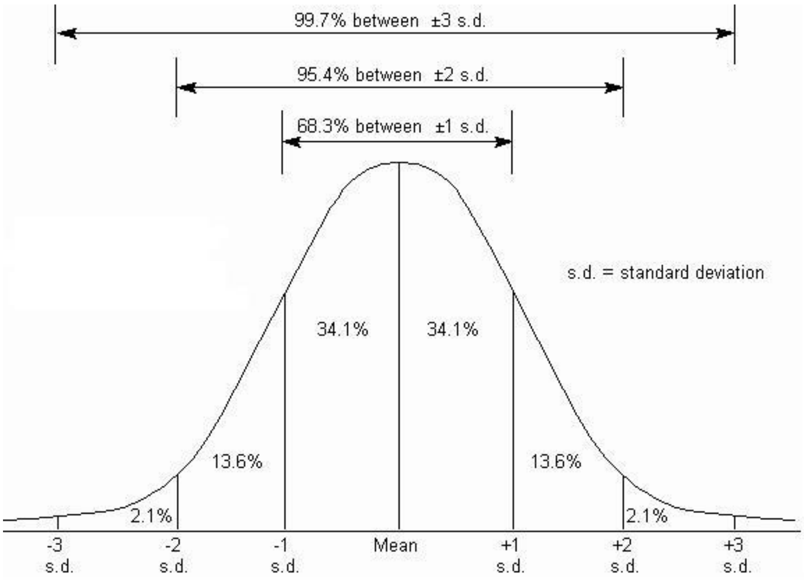
SD = Standard deviation

X = value

$\bar{x}$ = mean

n = the number of values you have

± 2 Standard Deviations



When data has a 'normal distribution' as shown in the lovely 'bell-shaped' graph above, we can see that **just over 95%** of the data is within two standard deviations either side of the mean. We can make use of this when comparing data.

± 2 Standard deviations   Worked Example 1. Lions

You have found the following ages of lions in two different zoos. The lions were randomly selected from all the lions in each zoo.

Age of Lions at Bristol Zoo (months)	Age of Lions at London Zoo (months)
36	46
31	50
35	48
24	49
21	51
47	49

You can use the calculators to calculate the mean and standard deviation, using the instructions on the calculator to help!

The standard deviation is calculated using the formula:

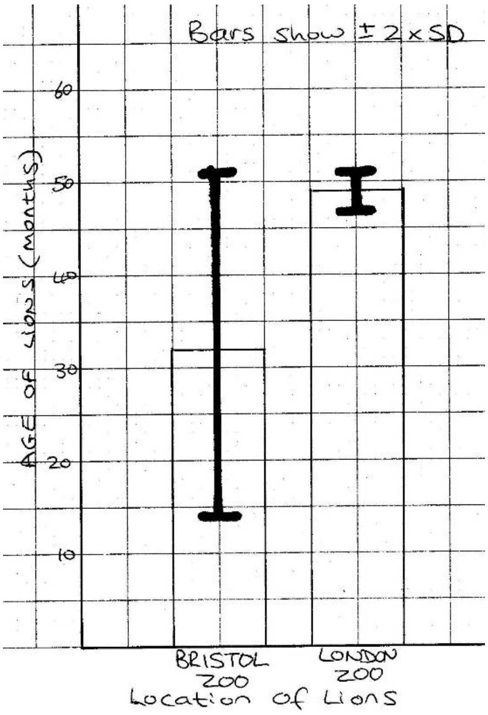
$$SD = \sqrt{\frac{\boldsymbol{\Sigma}(\mathbf{x} - \bar{\mathbf{x}})^2}{\mathbf{n} - \mathbf{1}}}$$

The best way to show these calculations is in do this is in a table.

	Bristol Zoo	**London Zoo**
mean	32.3	48.8
SD	9.3	1.7
2 x SD	18.6	2.4
Mean + (2 x SD)	50.9	51.2
Mean - (2 x SD)	13.7	46.4

Describing the results

We can draw a bar chart of the mean and plot the ± 2 Standard deviations from the mean and look at the overlap of the bars.

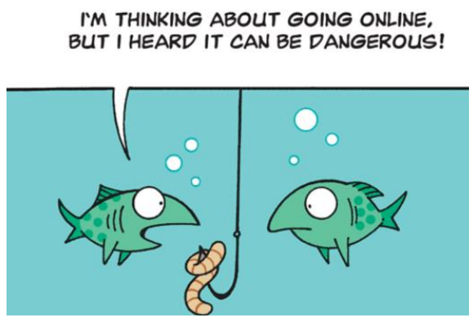


There is an overlap in the (±2 SD) bars.

This indicates that the differences in the means (the age of the Lions at Bristol zoo and London zoo) are **likely** to be due to chance.

Note: You cannot say how 'likely' this is due to chance – just that it is likely!

± 2 Standard deviations Worked Example 2.Fish



You have measured the sizes of the different genders of a species of tropical fish.

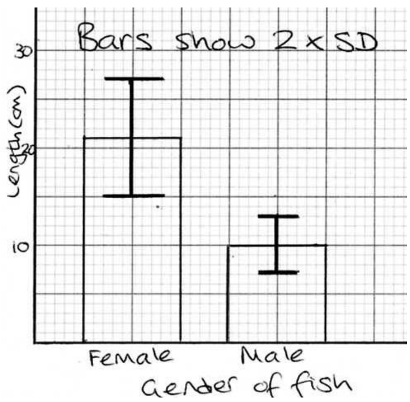
Length of female fish (cm)	Length of male fish (cm)
18	8
18	10
21	10
23	12
25	

Now calculate the mean and standard deviation.

	Female	**Male**
mean	21	10
SD	3	1.6
2 x SD	6	3.2
Mean + (2 x SD)	27	13
Mean - (2 x SD)	15	7

Describing the results

We can draw a bar chart of the mean and plot the ± 2 Standard deviations from the mean and look at the overlap of the bars.

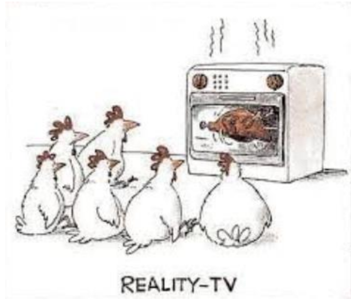


There is no overlap in the (±2 SD) bars.

This indicates that the differences in the means (the size of the fish) is **unlikely** to be due to chance.

Note: You cannot say how 'unlikely' this is due to chance – just that it is unlikely!

± 2 Standard deviations Question 1. Heart Rate



Compare the data for resting heart rate whilst watching two different TV shows. Describe the data.

Heart rate (beats per min) whilst watching ..	
Judge Judy	**Judge Rinder**
117	95
156	155
124	131
128	160
139	145
143	98

Now calculate the mean and standard deviation.

mean		
SD		
2 x SD		
Mean + (2 x SD)		
Mean - (2 x SD)		

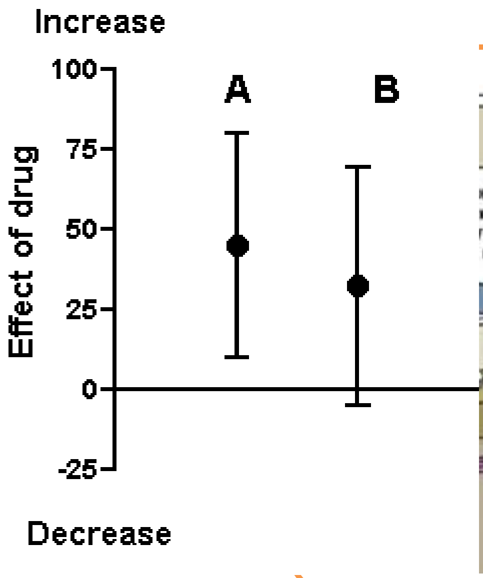
Describing the results

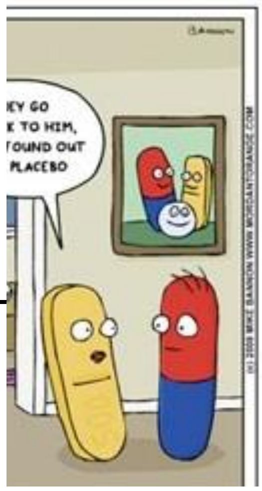
Plot the bar chart on graph paper and draw on the 2 x SD bars. You may find you can describe the data without plotting the bar chart though.

There is **an / no** overlap in the (±2 SD) bars.

This indicates that the differences in the means (……………………………..) is **unlikely/likely** to be due to chance.

± 2 Standard deviations Question 2. Drugs





Describing the results

Compare the data shown in the graph above. This figure depicts two experiments, A and B. In each experiment, control and treatment measurements were obtained. The graph shows the mean difference between control and treatment for each experiment. A positive number denotes an increase; a negative number denotes a decrease. The bars show the 2 x SD for those differences.

.

There is **………………………………………………**

This …………………………………………………………………………………………

…………………………………………………………………………………………………

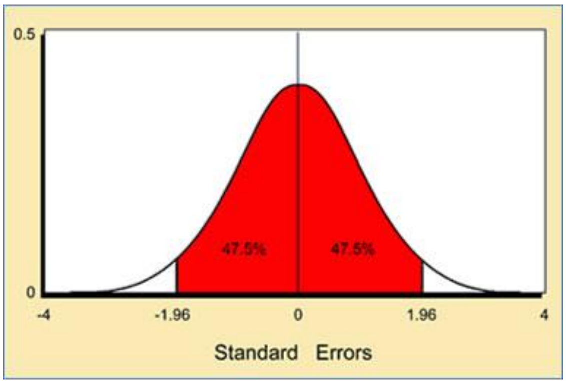
Standard Error of the Mean and 95% confidence interval

The **Standard Error** of a mean is calculated using the following formula:

$$SE = \frac{SD}{\sqrt{n}}$$

`

When you take a sample and calculate the mean, it is important to remember that this is only an estimate of the true mean for the whole of the population you are measuring. If you took a second sample, you would probably arrive at a slightly different estimate of the mean. There is no reason to suppose that the real mean will be exactly equal to the sample mean. It is likely to be close to it, however, and the amount by which it is likely to differ from the estimate can be found from the standard error.



What we do is find values (either side of the sample mean) which are likely to include the real mean, and say that we estimate the real mean to lie somewhere between the upper and lower values. If you look at the graph you can see that there is a 95% chance that the true mean is within 1.96 x SE either side of the mean of your sample.

The **95% Confidence Interval** is then calculated using the following formula:

$$\mathbf{95\%}\ \boldsymbol{CI} = \bar{\mathbf{x}} \pm \mathbf{SE} \times \mathbf{1.96}$$

We can use the 95% confidence interval to state that:

- we are 95% confident that the true mean value of the population from which the sample was taken lies between the upper and lower confidence limits
- if the intervals of two calculated means do not overlap, we are 95% confident that these means are different.

NOTE: People find the terms ‘standard error’ and ‘standard deviation’ confusing. We use the term‘standard deviation’ when we are talking about distributions, either of a sample or a
population. We use the term ‘standard error’ when we are talking about an estimate found from a sample. If we want to say how good our estimate of the mean measurement is, we quote the standard error of the mean. If we want to say how widely scattered the measurements are, we quote the standard deviation.

Standard error of the Mean and 95% Confidence Interval

Worked Example 1. The mean mass of Guinea Pigs

The owner of a Pet shop wants to work out the average mass of small guinea pigs in her shop. She only had time to measure the mass of a few which she randomly caught. Calculate the mean mass with 95% confidence intervals, using the sample data below.

Mass of guinea pigs (g): 19.4 21.4 22.3 22.1 20.1 23.8 24.6 19.9 19.1

Now layout the calculations (in a table helps)

mean	21.4
n	9
√n	3
SD	1.95
SE (SD ÷ √n)	0.65
1.96 x SE	1.27
Mean + 1.96 x SE	22.7
Mean - 1.96 x SE	20.1

Describing the results

The mean mass (of the **sample** of Guinea pigs) = 21.4g

We are **95% confident** that the true mean value

(of the **whole** population of Guinea pigs in the pet shop)

is between 20.1g – 22.7g

Standard error of the mean and 95% confidence interval

Worked Example 2. Limpets

A resident on the Isle of Wight wants to know if the size of limpets is different between the upper and middle ledges at Bembridge rocky shore. The resident collects the following data. (Obviously the resident is not a proper scientist otherwise they would never have presented the data in such a terrible way!)   (Sizes in mm, measured using callipers:)

Middle: 43.9, 38.4, 39.4, 44.7, 40.3, 37.8, 35.6, 56.7, 47.3, 36.0, 38.7, 37.7, 41.6, 42.5, 48.3, 34.9, 21.3, 42.5, 36.2, 46.8, 41.6, 43.5, 48.6, 39.5, 50.2, 46.9, 48.7, 42.9, 37.6, 53.4

Upper: 40.3, 35.0, 36.2, 40.8, 31.2, 30.2, 27.3, 37.7, 26.5, 25.5, 33.6, 24.8, 42.3, 38.6, 30.7, 23.2, 37.3, 34.6, 32.3, 33.0, 34.6, 38.0, 28.3, 22.5, 37.0, 45.0, 32.8, 36.9, 44.2, 28.6

The best way to show these calculations is in a table.

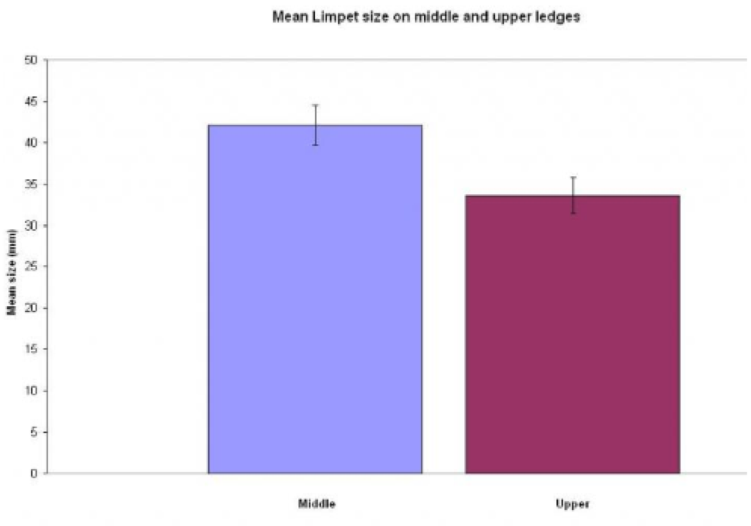
	Middle ledge	**Upper ledge**
mean	42.12	33.63
n	30	30
√n	5.48	5.48
SD	6.74	6.07
SE (SD ÷ √n)	1.23	1.11
1.96 x SE	2.41	2.18
Mean + 1.96 x SE	44.53	35.81
Mean - 1.96 x SE	39.71	31.45

Interpreting the results

The 95% CI of the size of snails on the Middle ledge are 39.71mm to 44.53mm.

The 95% CI of the size of snails on the Upper ledge are 31.45mm to 35.81mm.

There is no overlap. Some people prefer to use graph paper to draw a bar chart with bars to show the confidence intervals, but it is not essential to do this. The 95% CI bars do not overlap.



There is no overlap in the 95% confidence intervals of the two calculated means (the mean size of Limpet on the Middle ledge and Limpets on the Upper ledge).

We can be 95% confident that the means are different.

Standard error of the mean and 95% confidence interval

Worked Example 2. Weight

In the UK almost a third of adults are obese. A new diet called the South Beach diet has become popular, and the NHS wants to assess how effective it is compared to a more traditional low calorie diet. The following data is available. What conclusions should be drawn from the available data?

% Reduction in BMI (body mass index) after 4 weeks of completion of diet	
South Beach diet	**Traditional Low Calorie diet**
2.1	2.2
2.0	1.8
1.8	1.9
2.0	2.0
1.9	1.9
2.4	

There are six values for the South Beach diet so n= 6, and five values for the traditional low calorie diet so n=5.

	South Beach diet	Traditional Low Calorie diet
mean	2.0	2.0
n	6	5
$\sqrt{n}$	2.45	2.24
SD	0.21	0.15
SE (SD ÷ $\sqrt{n}$)	0.08	0.07
1.96 x SE	0.16	0.14
Mean + 1.96 x SE	2.2	2.1
Mean - 1.96 x SE	1.8	1.9

Interpreting the results

The 95% confidence intervals for the South Beach diet is 1.8 to 2.2 percent.

The 95% confidence intervals for the South Beach diet is 1.9 to 2.1 percent.

There is an overlap in the 95% confidence intervals of the two calculated means (of the % reduction in BMI for the South beach diet and the traditional low calorie diet).

We can be 95% confident that the means are not different.

Standard error of the mean and 95% Confidence Interval

Question 1. The mean abundance of green-winged orchid

(Use Worked example 1 to help with this question)

The counts of the green-winged orchid in a sample of ten 1x1m quadrats placed randomly in a hay meadow are shown below:

4 2 58 1 22 5 7 17 1 4

Calculate the mean count of the orchid (per metre^2) with 95% confidence intervals.

mean	
n	
√n	
SD	
SE (SD ÷ √n)	
1.96 x SE	
Mean + 1.96 x SE	
Mean - 1.96 x SE	

Describing the results

The mean count of the sample of orchids per metre^2 = …………..

We are **…………………………..** that the true mean value

(of the **whole** population of …………………………………)

is between …………………………………..

Standard error of the mean and 95% confidence interval

Question 2. Beans

Q: What's the fastest vegetable?

A: A runner bean

Two students have each grown a row of runner beans in the school allotment. Each day Paul has talked to his plants to encourage them to grow taller. Simon has not, and he thinks that Paul is a bit bonkers. Is there any evidence to suggest that talking to the plants has made them grow taller? (Use worked example 2 & 3 to help with these questions.)

Height of Runner bean plants (cm)	
Paul's plants	**Simon's plants**
124	131
156	155
128	160
139	145
117	95
142	65
123	117
98	212
153	160
120	60

Calculating Standard error and 95% confidence interval

	Paul's plants	**Simon's plants**
mean		
n		
√n		
SD		
SE (SD ÷ √n)		
1.96 x SE		
Mean + 1.96 x SE		
Mean - 1.96 x SE		

Interpreting the results

There is **an/no** overlap in the ………………………………. of the two calculated means (of the mean height of Paul or Simon's bean plants).

We can be 95% confident that the means **are/are not** different.

Standard error of the mean and 95% confidence intervals

Q3. More Beans

Q: What kind of vegetable is jealous? A: A green bean!

Ultra-competitive Paul was a bit disappointed at the result of his last experiment, so this time Paul has decided to grow his beans using a super new fertiliser, whereas Simon just likes to use horse-manure. Can a conclusion be made about which fertiliser is best?

Height of Runner bean plants (cm)	
Paul's plants (grown using new fertiliser)	**Simon's plants (grown using horse manure)**
103	160
113	145
117	120
123	125
119	126
98	212
124	123
120	129

Calculating Standard error and 95% confidence interval

Interpreting the results

There ………

We ……

More Standard error and 95% confidence interval problems.

4. Spotted knapweed is a common weed in the USA. Two methods, chemical control and biological control, have been used to reduce the numbers of spotted knapweed plants. Use Standard error and 95% confidence intervals to determine whether biological control is better than biological control?

Mean number of spotted knapweed plants per m^2	
Chemical control	**Biological control**
2	2
15	3
3	3
20	5
3	4
16	3
2	2

5. Two fields, A and B, were used to grow the same crop. Random samples of crop plants from each plot were collected and their mass determined. The results are shown in the table. Does the evidence suggest that previous use of the field affects the mass of the crop?

Sample	**Mass of crop/kg m^{-2}**	
	Field **A** - used for grazing cattle in previous year	Field **B** - used for same crop in previous year
1	14.5	6.4
2	16.7	9.8
3	17.4	12.9
4	17.5	16.2
5	17.5	17.1
6	17.5	17.1
7	17.5	17.1

6. Mayflies are insects which lay their eggs in streams and rivers. The nymphs hatch from the eggs and live in the water for several years. Mayfly nymphs were collected by disturbing the gravel of a stream bed. A net placed immediately downstream caught any animals which were washed out of the gravel. Eight samples were collected from shallow, fast-flowing parts of the stream and eight from deeper, slow-flowing parts. The results are given in the table. What can be concluded about the differences in the numbers of nymphs found in shallow water and deep water?

Sample Number	Number of Mayfly nymphs found	
	Shallow water	Deep Water
1	12	16
2	14	14
3	13	14
4	18	18
5	16	21
6	16	17
7	13	16
8	14	15

7. The table shows the relative thickness of the walls of the aorta and vena cava in four different people. Is it possible to show that the thickness of arteries is thicker than veins?

Name	**Thickness / μm**	
	Artery	**Vein**
John	420	180
James	490	175
Daniel	370	165
David	120	120

8. A general practitioner has been investigating whether the diastolic blood pressure of men aged 20-44 differs between farm workers and firemen. For this purpose, she has obtained a random sample of 72 firemen and 48 farm workers and calculated the mean and standard deviations. What conclusions can be made?

	Farm workers	**Firemen**
Mean diastolic blood pressures (mmHg)	88	79
Standard Deviation	4.5	4.2

Using what we know to make inferences about what we don’t know

Inferential Statistics

You should be aware that inferential statistics are used to test a theory, known as a hypothesis. Before we choose a statistical test we should write a null hypothesis. The table shows how hypotheses can be turned into null hypotheses.

Hypothesis	**Null hypothesis**
Chickens fed maize supplemented by lipid produce more male offspring than those fed maize alone.	There is no difference between the number of male and female offspring of chickens fed maize supplemented by lipid and those fed maize alone.
There are fewer slugs in dry areas	There is no difference between the number of slugs found in wet and dry areas
Tobacco plants exhibit a higher rate of growth when planted in soil rather than peat	Tobacco plants do not exhibit a higher rate of growth when planted in soil rather than peat.

The null hypothesis can be thought of as “life is dull and nothing ever happens” hypothesis!

Once the null hypothesis is stated, a statistical test is then chosen. This either supports or fails to support the null hypothesis.

Chi-squared test

All chi-squared tests are concerned with counts of things (frequencies) that you can put into categories. For example, you might be investigating flower colour and have counted the numbers (frequencies) of red flowers and white flowers (categories). Or you might be investigating human health and have frequencies of smokers and non-smokers.

The test looks at the frequencies you obtained when you counted them and compares them with the frequencies you might **expect to get** in order to determine whether the difference is significant or not.

The χ^2 test

The chi-square (χ^2) test is based on calculating the value of χ^2 from the equation

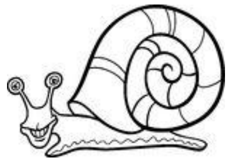
$$\chi^2 = \sum\frac{(O-E)^2}{E}$$

where O represents the results you observe in the investigation
and E represents the results you expect.

Table showing the critical values of χ^2 at P = 0.05 for different degrees of freedom

Degrees of Freedom	Critical value
1	3.84
2	5.99
3	7.82
4	9.49
5	11.07
6	12.59
7	14.07
8	15.51
9	16.92
10	18.31

Chi-squared Worked Example 1. Snails on the Seashore



Why did the periwinkle blush?

Answer: because the sea weed!!

You have been wandering about on a seashore and you have noticed that a small snail (the flat periwinkle) seems to live only on certain types of seaweed. You decide to investigate whether the animals prefer to certain types of seaweed by counting numbers of animals on the different types of seaweed. You end up with the following data:

Type of Seaweed	**Observed frequency** **(the numbers of periwinkle)**
serrated wrack	45
bladder wrack	38
egg wrack	10
spiral wrack	5
other seaweed	2
TOTAL	100

Null hypothesis

The null hypothesis when doing Chi-squared is

“**there is no significant difference between the observed and expected frequencies**.”

In other words, the periwinkle does not have a preference for which seaweed it lives on. This is now used to work out the ‘expected’ frequencies.

Expected Frequencies

Our null hypothesis is that there is no difference between the observed and expected frequencies. If this were exactly the case there would be no differences in the frequencies over all of our categories (i.e. the five types of seaweed). The best estimate we could make therefore would be to add up all our observed frequencies and divide by the number of categories. So our expected frequency for each category would be:

45 + 38 + 10 + 5 + 2 = 100

100 ÷ 5 = 20 = expected frequency

Type of Seaweed	Observed frequency	Expected frequency
serrated wrack	45	20
bladder wrack	38	20
egg wrack	10	20
spiral wrack	5	20
other seaweed	2	20
TOTAL	100	100

Calculating the Chi-squared value

Next we calculate the value of Chi-squared using the formula below.

$$\chi^2 = \sum\frac{(O-E)^2}{E}$$

O = Observed frequency

E = Expected frequency

The best way to show these calculations is in do this is in a table.

Type of Seaweed	Observed frequency	Expected frequency	O-E	$(O-E)^2$	$\frac{(O-E)^2}{E}$
serrated wrack	45	20	25	625	31
bladder wrack	38	20	18	324	16
egg wrack	10	20	-10	100	5
spiral wrack	5	20	-15	225	11
other seaweed	2	20	-18	324	16
					Total = 79

The total of our final column is the Chi-squared value. (χ^2 = 79)

The ‘critical value’ & the ‘degrees of freedom’

Before we can interpret our results we need to work out the ‘critical value’. The critical value represents the borderline between accepting or rejecting our null hypothesis. We get the critical value from the data sheet, but this depends on the number of ‘degrees of freedom’.

Table showing the critical values of χ^2 at P = 0.05 for different degrees of freedom

Degrees of Freedom	Critical value
1	3.84
2	5.99
3	7.82
4	9.49
5	11.07
6	12.59
7	14.07
8	15.51
9	16.92
10	18.31

Degrees of freedom = number of categories -1

'Degrees of freedom' is a term that can be bit confusing. A simple (though not completely accurate) way of thinking about degrees of freedom is to imagine you are picking people to play in a team. You have eleven positions to fill and eleven people to put into those positions. How many decisions do you have? In fact you have ten, because when you come to the eleventh person, there is only one person and one position, so you have no choice. You thus have ten 'degrees of freedom' as it is called. So 11 categories but only 10 ‘degrees of freedom’. Hence, **degrees of freedom = number of categories -1**

Likewise, the periwinkle snail was found on the serrated wrack, bladder wrack, egg wrack, spiral wrack, or other seaweed. There are five categories (five different types of seaweed), so only **4 degrees of freedom**.

Interpreting the results

Chi-squared gives a number which indicates how big the difference is between the observed data and the expected data.

If the Chi-squared value is small, then there is a small difference between the observed and the expected data. This means the null hypothesis is accepted (likely to be correct). In other words, the snails don’t mind which seaweed they live on!

If the Chi-squared value is huge, then there is a huge difference between the observed and the expected data. This means the null hypothesis is rejected. In other words, the snails do indeed have a preference for living on a particular seaweed.

Looking at the table above we can see that the critical value of Chi-squared at 5% significance (p=0.05) and 4 degrees of freedom is **9.49**.

Our calculated value is **79**

The calculated value is bigger (much bigger!) than the critical value. In a chi-squared test this means we must reject the null hypothesis. In doing this we are saying that the snails are not scattered about the various sorts of seaweed randomly. Biologists would infer that this means they seem to prefer living on certain species.

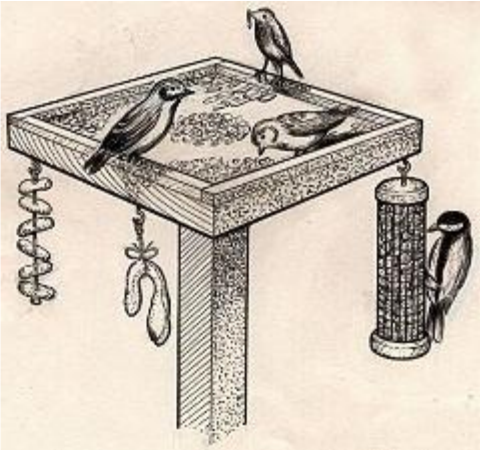
Our calculated value of Chi-squared is much larger than the critical value of Chi-squared.

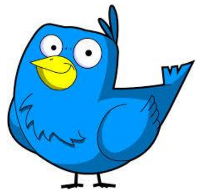
There is less than 5% probability that the differences (between the observed and expected data) are due to chance.

We reject our null hypothesis.

It is worth pointing out that statistics of this kind tell you nothing about the biology of the situation. All we are saying is that our observed frequencies are different to our expected ones. (For example, you could criticise our approach by pointing out that it might be that there are not equal amounts of each type of seaweed on the shore for the animals to live on.)

Chi-squared Worked Example 2. Birds on the Bird-Table





Q: When should you buy a bird?
A: When it's going cheep!

Three neighbours have very similar bird-tables in their gardens. Bill owns the middle garden and is a keen birdwatcher but he suspects that Kate, one of his neighbours, is actively encouraging birds away from Bill's garden into her own somehow. Is Bill right?

Garden	**Observed frequency** **(the numbers of bird's visiting the garden on one day in March)**
Bill	112
Kate	145
Chris	139
TOTAL	396

Null hypothesis

The null hypothesis when doing Chi-squared is

"**there is no significant difference between the observed and expected frequencies**."

Expected Frequencies and Calculating the Chi-squared value

We would expect the same numbers of birds to visit each garden if the null hypothesis is correct. A total of 396 birds were seen, so we would expect 132 in each garden (396 ÷ 3)

Garden	**Observed frequency**	**Expected frequency**	**O-E**	**$(O-E)^2$**	**$\frac{(O-E)^2}{E}$**
Bill	112	132	-20	400	3.0
Kate	145	132	13	169	1.3
Chris	139	132	7	49	0.4
					χ^2 **= 4.7**

Degrees of Freedom

Degrees of freedom = number of categories -1

We have three categories (i.e. three gardens) so the degrees of freedom is 3-1 = **2**

Interpreting the results

Looking at the table above we can see that the critical value of Chi-squared at 5% significance (p=0.05) and 2 degrees of freedom is **5.99.**

Our calculated χ^2 value is **4.7**

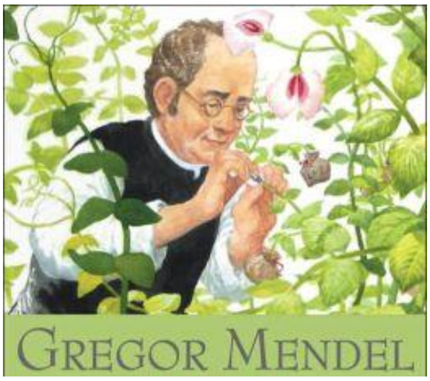
The calculated value is smaller than the critical value. In a chi-squared test this means we must accept the null hypothesis. (In other words, Bill is not right! – Kate is **not** secretly putting out expensive bird-food to encourage the birds into her garden!)

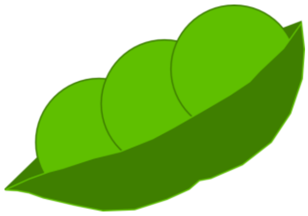
Our calculated value of Chi-squared is smaller than the critical value of Chi-squared.

There is more than 5% probability that the differences (between the observed and expected data) are due to chance.

We accept our null hypothesis.

Chi-squared Question 1. Mendel and his Peas





Q: What do you call an angry pea?
A: Grump-pea.

Mendel planted some round peas which grew into plants that produced a total of 556 peas. 423 were round peas and 133 were wrinkled peas. Mendel postulated that round is dominant to wrinkled, and he work out the expected ratio for two heterozygous parent plants as 3:1. Do his experimental data support his 3:1 expected ratio?

Null hypothesis Write the null hypothesis here.

……………………………………………………………………………………………………..

……………………………………………………………………………………………………..

Calculating the Chi-squared value

Shape of pea	Observed frequency	Expected frequency	O-E	$(O-E)^2$	$\frac{(O-E)^2}{E}$
					$\chi^2 =$

Interpreting the results

How many degrees of freedom are there? ………………

What is the critical value (from the table)? …………………..

Our calculated value of Chi-squared is larger/smaller than the critical value of Chi-squared.

There is more/less than 5% probability that the differences (between the observed and expected data) are due to chance.

We accept/reject our null hypothesis.

Chi-squared Question 2. A zookeeper's dilemma

Q: What do you call a flying primate?   A: A hot air **baboon**!

A zookeeper thinks that lowering the intensity of the light in the primate exhibits will reduce the amount of aggression between the baboons. In exhibit A, with a lower light intensity, he observes 36 incidences of aggression over a one month period. In exhibit B, with normal lights, he observes 42 incidences of aggression. Does he have enough evidence to support his theory?

Null hypothesis   Write the null hypothesis here.

…………………………………………………………………………………………………..

…………………………………………………………………………………………………..

Calculating the Chi-squared value

Interpreting the results

How many degrees of freedom are there? ………………

What is the critical value (from the table)? …………………..

Our calculated value of Chi-squared is larger/smaller than the critical value of Chi-squared.

There is more/less than 5% probability that the differences (between the observed and expected data) are due to chance.

We accept/reject our null hypothesis.

Chi-squared Question 3. The West-African bee-eaters

You have just returned from a 3 year stint in the jungles of western Africa, where you studied the habitat selected by the native bee-eaters (a family of birds that specialize in catching bees and wasps on the wing, taking them to a perch, bashing their stingers out, and devouring them. In a pinch, they will eat other flying or hopping insects, such as grasshoppers). Several habitats were available to the bee-eaters. Is there evidence to suggest that birds prefer a particular habitat?

	Forest floor	Understory layer	Canopy layer	Emergent layer	Grassland	Field	River-bank
Number of birds	3	15	17	20	3	11	4

Null hypothesis Write the null hypothesis here.

……………………………………………………………………………………………………..

……………………………………………………………………………………………………..

Calculating the Chi-squared value

Interpreting the results

How many degrees of freedom are there? ………………

What is the critical value (from the table)? …………………..

Our calculated value of Chi-squared is larger/smaller than the critical value of Chi-squared.

There is more/less than 5% probability that the differences (between the observed and expected data) are due to chance.

We accept/reject our null hypothesis.

More Chi-squared problems.

4. One section of a river was trawled and four species of fish counted and frequencies recorded. There were 15 Rudd, 15 Roach, 4 Dace and 6 Bream. Are the fish present in the river in equal proportions?

5. An optician noticed the following information about colour-blindness in males and females. Is there a significant difference between the between the observed frequency of colour blindness in males and females?

Observed frequencies	Males	Females
Colour blind	56	14
Not colour blind	754	536

6. The table below shows the number of patients requesting an urgent appointment to see a Doctor on particular days of the week. Are the differences significant?

	Monday	Tuesday	Wednesday	Thursday	Friday
Numbers of patients	125	88	87	94	108

7. Cranes are large birds. Biologists have used DNA hybridisation to confirm the relationships between different species of crane. They made samples of hybrid DNA from the same and from different species. They measured the percentage of hybridisation of each sample. The results are shown in the table. Are there any differences which are statistically significant?

Species of crane	**Mean percentage DNA hybridisation**
Grus americana and *Grus monachus*	97.4
Grus monachus and *Grus rubicunda*	95.7
Grus americana and *Grus rubicunda*	95.5
Grus rubicunda and *Grus rubicunda*	99.9
Grus americana and *Grus americana*	99.9
Grus monachus and *Grus monachus*	99.8

8. The table shows the number of cases of tuberculosis in the East Midlands between 2000 and 2005. Are the differences in number of cases of tuberculosis significant?

Number of cases of TB per 100,000 of the population	2000	2001	2002	2003	2004
	10.6	11.1	11.9	7.9	9.9

Spearman’s Rank

Spearman rank correlation test

Calculate the value of the Spearman rank correlation, r_s, from the equation

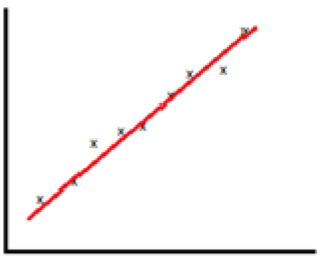
$$r_s = 1 - \left[\frac{6 \times \Sigma D^2}{n^3 - n}\right]$$

where n is the number of pairs of items in the sample and D is the difference between each pair of ranked measurements.

Table showing the critical values of r_s at P = 0.05 for different numbers of paired values

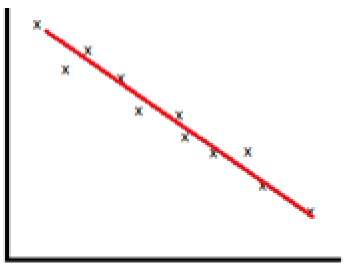
Number of pairs of measurements	Critical value
5	1.00
6	0.89
7	0.79
8	0.74
9	0.68
10	0.65
12	0.59
14	0.54
16	0.51
18	0.48

Spearman’s rank correlation is a statistical test that is carried out in order to assess the degree of association between different measurements from the same sample. That is, if you are looking for a positive or negative correlation between two variables.



Positive correlation

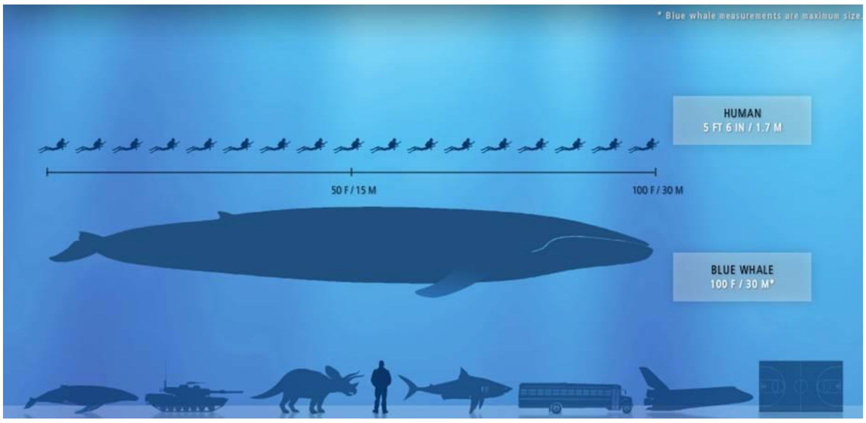
No correlation



Negative correlation

When the points on a graph clearly fit onto a line of best fit it is easy to determine whether a correlation exists. However, as the points become further placed from each other it is hard to make an accurate judgement. This is where statistics is used; to clarify how confident we are that a correlation exists.

Spearman’s Rank Worked Example 1. Whales



Q: What do you get if you cross an elephant with a whale?

A: A submarine with a built-in snorkel.

Do blue whales get heavier as they get longer? From the table below it certainly looks as if they do. If we drew a graph we would get an 'uphill line' (a positive correlation). However, to find out if the correlation is statistically significant we must calculate Spearman's rank correlation coefficient (r_s).

Length (metres)	Mass (tonnes)
1	1.5
1.5	2.3
2.3	3.6
3.4	7.1
4.4	2.6
4.6	13.3
6.2	7.5
7	11.2
7	12.1
8.7	11
10.5	12
12	18.2

Null hypothesis

When doing Spearman's rank the two variables are used to construction the null hypothesis. The null hypothesis always assumes there is no relationship.

"There is no correlation between (variable 1) and (variable 2)"

In other words, for this example:

"There is no correlation between the length of a blue whale and the mass of the whale."

Rank the data

For each set of data assign ranks from lowest to highest. The lowest value in a column will be given the rank of 1, the next smallest number will be given a 2 etc. If there are tied scores each of those will share the ranks and be given the average (mean) rank.
For example, there are two whales with the same length (7 metres). These would be ranked 8 and 9, so they are given the mean value. (i.e. 8+9 = 17   17÷2 = 8.5).

Length metres) (variable 1)	Rank of variable 1	**Mass (tonnes)** (variable 2)	Rank of variable 2
1	1	1.5	1
1.5	2	2.3	2
2.3	3	3.6	4
3.4	4	7.1	5
4.4	5	2.6	3
4.6	6	13.3	11
6.2	7	7.5	6
7	8.5	11.2	8
7	8.5	12.1	10
8.7	10	11	7
10.5	11	12	9
12	12	18.2	12

Calculate Spearman's rank

Next we have to work out D^2 ;this is the difference in the rankings, squared.

Length/m	Rank length	Mass/tonnes	Rank mass	Difference/*D*	D^2
1.0	1	1.5	1	0	0
1.5	2	2.3	2	0	0
2.3	3	3.6	4	–1	1
3.4	4	7.1	5	–1	1
4.4	5	2.6	3	2	4
4.6	6	13.3	11	–5	25
6.2	7	7.5	6	1	1
7.0	8.5	11.2	8	0.5	0.25
7.0	8.5	12.1	10	–1.5	2.25
8.7	10	11.0	7	3	9
10.5	11	12.0	9	2	4
12.0	12	18.2	12	0	0
			Σ	0	47.5

Then we calculate the value of Spearman's rank correlation, r_s, using the equation below.

$$r_s = 1 - \left[\frac{6 \times \sum D^2}{n^3 - n}\right]$$

D = difference between the rank of the paired measurements

n = number of paired measurements

∑ = the sum of

Add up the D^2 column ($\sum D^2$). $\sum D^2$ = **47.5**

If we substitute the numbers into the equation we get:

$$r_s = 1 - \left[\frac{6 \times \sum D^2}{n^3 - n}\right]$$

$= 1- ((6 \text{ X } 47.5) \div (12^3\text{-}12)$

$= 1 - (285 \div 1716)$

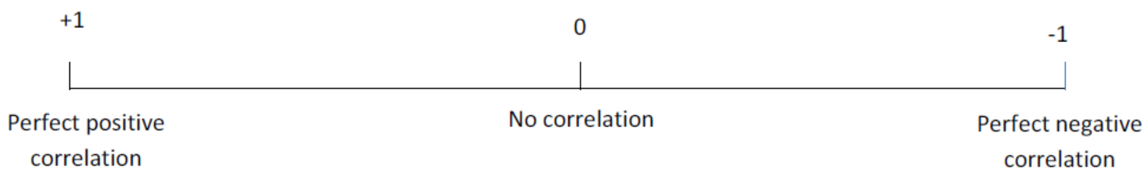
$= 1 – 0.166$

r_s = 0.834

<u>Interpret the results</u>

The first thing we notice is that the answer is a positive number, so we know that length are positively correlated (i.e. not negatively correlated.)

The closer r_s is to 1 or -1 the more likely the correlation. A perfect positive correlation has an r_s value of 1, a perfect negative correlation has a value of -1.



If the value lies between -1 and 1 we need to carry out a test for significance.

Before we can interpret our results we need to work out the 'critical value'. The critical value represents the borderline between accepting or rejecting our null hypothesis.

Table showing the critical values of r_s at P = 0.05 for different numbers of paired values

Number of pairs of measurements	Critical value
5	1.00
6	0.89
7	0.79
8	0.74
9	0.68
10	0.65
12	0.59
14	0.54
16	0.51
18	0.48

We have 12 paired values which gives us a critical value of 0.59 (for a positive correlation) or -0.59 (for a negative correlation). This means that any value between +0.59 and +1 is a statistically significant (positive) correlation.

The Spearman's rank correlation coefficient, r_s, between the length and mass of a blue whale was **+0.834**. The calculated value is bigger than the critical value. In a Spearman's rank test this means we must reject the null hypothesis. In doing this we are saying that there is a relationship (a positive correlation) between the length and mass of the blue whale.

Our calculated value of Spearman's rank correlation coefficient, r_s (+0.834) is larger than the critical value of (+0.59).

There is less than 5% probability that the positive correlation between the length and mass of the whale is due to chance.

We reject our null hypothesis.

Spearman's Rank Worked Example 2. Heather in a Moorland

Q: What did the big flower say to the small flower?
A: What's up, bud?

During monitoring of a moor, an ecologist collected data from an area of moorland that had been restored in 2003. In order to assess whether a correlation existed between the two plant species she studied or whether they were growing independently of one another, she used a quadrat to collect the following data.

Number of Bilberry plants (per m^2)	Number of Common Heather plants (per m^2)
50	18
175	12
270	20
375	10
425	10
580	12
710	8
790	6
890	10
980	9

Null hypothesis

"There is no correlation between the number of Bilberry plants growing in an area and the number of Common Heather plants growing in the area"

Rank the data & Calculate Spearman's rank

Number of Bilberry plants (per m^2)	Rank Bilberry (variable 1)	Number of Common Heather plants (per m^2)	Rank Common Heather (variable 2)	Difference (D)	D^2
50	1	18	9	-8	64
175	2	12	7.5	-5.5	30.25
270	3	20	10	-7	49
375	4	10	5	-1	1
425	5	10	5	0	0
580	6	12	7.5	-1.5	2.25
710	7	8	2	5	25
790	8	6	1	7	49
890	9	10	5	4	16
980	10	9	3	7	49
					$\sum D^2$ = 285.5

$\sum D^2$ = **285.5**

n = 10 (because we have 10 paired readings for the plants)

If we substitute the numbers into the equation we get:

$$r_s = 1 - \left[\frac{6 \times \sum D^2}{n^3 - n}\right]$$

= 1- ((6 X 285.5) ÷ (10^3-10)

= 1 - (1713 ÷990)

= 1 – 1.730

r_s = - 0.730

Interpret the results

The answer is a negative number, so we know that relationship between the two plants shows a negative correlation. We have 10 paired values which gives us a critical value (from the table of critical values) of -0.65 (for a negative correlation).

Our calculated value of Spearman's rank correlation coefficient (-0.730) is larger than the critical value of (-0.65).

There is less than 5% probability that the negative correlation between the Bilberry and Common Heather is due to chance.

We reject our null hypothesis.

Spearman's Rank Worked Example 3. Frigate birds

Males of the Frigatebird have a large red throat pouch. They visually display this pouch and use it to make a drumming sound when seeking mates. Researchers wanted to know whether females, who presumably choose mates based on their pouch size, could use the pitch of the drumming sound as an indicator of pouch size. They estimated the volume of the pouch and the frequency of the drumming sound in 16 males:

Volume (cm^3)	Frequency (Hz)
1760	529
2040	390
2440	473
2550	461
2730	465
2740	532
3010	484
3080	527
3370	488
3740	485
4910	478
5090	434
5380	468
5850	449
6730	464
6990	530

Null hypothesis

"There is no correlation between the volume of the bird's pouch and frequency of the sound it makes."

Rank the data & Calculate Spearman's rank

Volume (per cm^3)	Rank Volume (variable 1)	Frequency (Hz)	Rank Frequency (variable 2)	Difference (D)	D^2
1760	1	529	15	-14	196
2040	2	390	1	1	1
2440	3	473	9	-6	36
2550	4	461	5	-1	1
2730	5	465	7	-2	4
2740	6	532	17	-11	121
3010	7	484	11	-4	16
3080	8	527	14	-6	36
3370	9	488	13	-4	16
3740	10	485	12	-2	4
4910	11	478	10	1	1
5090	12	434	3	9	81
5380	13	468	8	5	25
5850	14	449	4	10	100
6730	15	464	6	9	81
6990	16	530	16	0	0
					$\sum D^2 = 719$

$\sum D^2$ = **719**

n = 16

If we substitute the numbers into the equation we get:

$$r_s = 1 - \left[\frac{6 \times \Sigma D^2}{n^3 - n}\right]$$

= 1- ((6 X 719) ÷ (16^3-16)

= 1 - (4314 ÷ 4080)

= 1 – 1.057

r_s = - 0.057

<u>Interpret the results</u>

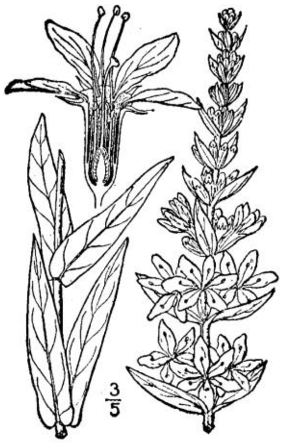
The answer is a negative number, so we know that any relationship between volume and frequency shows a negative correlation. We have 16 paired values which gives us a critical value (from the table of critical values) of -0.51 (for a negative correlation).

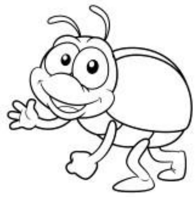
Our calculated value of Spearman's rank correlation coefficient (-0.057) is smaller than the critical value of (-0.51).

There is more than 5% probability that the correlation between the volume of the pouch and the frequency of the sound it makes is due to chance.

We accept our null hypothesis.

Spearman's rank Question 1. Purple loosestrife control





Q: What do you call a beetle that can't have too much sugar?

A: a diabeetle.

A European beetle was tested to see whether it could be used for the biological control of purple loosestrife in the USA. In an investigation, beetles were released in an area where purple loosestrife was a pest. The table shows some of the results. Is it possible to prove that the beetles are effective in controlling purple loosestrife?

Mean number of Purple loosestrife plants (per m^2)	Mean number of Beetles (per m^2)
28	4
22	5
8	40
6	62
7	68

Null hypothesis Write the null hypothesis here.

……..

……..

Rank the data & Calculate Spearman's rank

Mean number purple loosestrife (per m^2)	Rank Purple loosestrife (variable 1)	Mean number Beetles (per m^2)	Rank Beetles (variable 2)	Difference (D)	D^2
					$\sum D^2 =$

$\sum D^2$ = ...

n = ...

Substitute the numbers into the equation to calculate r_s

$$r_s = 1 - \left[\frac{6 \times \Sigma D^2}{n^3 - n}\right]$$

r_s = ...

<u>Interpret the results</u>

Our calculated value of Spearman's rank correlation coefficient is <u>smaller/bigger</u> than the critical value.

There is <u>less/more</u> than 5% probability that the <u>(positive/negative)</u> correlation between .... and .... is due to chance.

We <u>accept/reject</u> our null hypothesis.

Spearman's rank Question 2. Blood sugar control

Q: Did you hear the joke about the peanut butter?
A: I'm not telling you. You might spread it!

A student ate a meal containing carbohydrates at 07:00. He ate nothing else for the next five hours. The table shows the concentration of glucose in his blood at hourly intervals after the meal. Is there a significant relationship between time of day and blood sugar concentration?

Time of day	Concentration of glucose in blood (mg per 100 cm^3 of blood)
07:00	90
08:00	120
09:00	70
10:00	85
11:00	110
12:00	80

Null hypothesis Write the null hypothesis here.

……..

……..

Rank the data & Calculate Spearman's rank

					$\sum D^2$ =

$\sum D^2$ = ……………………

n = ………………………..

Substitute the numbers into the equation to calculate r_s

$$r_s = 1 - \left[\frac{6 \times \Sigma D^2}{n^3 - n}\right]$$

r_s = ………………

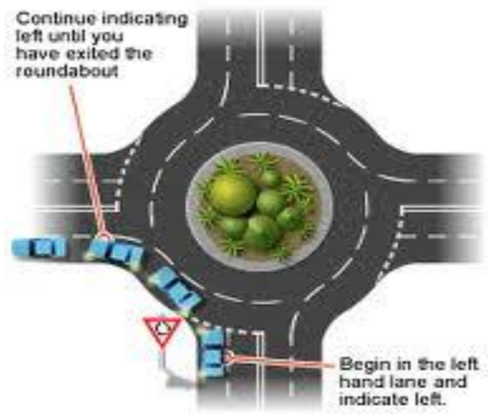
<u>Interpret the results</u>

Our calculated value of Spearman's rank correlation coefficient is <u>smaller/bigger</u> than the critical value.

There is <u>less/more</u> than 5% probability that the <u>(positive/negative)</u> correlation between ………………………………………………………………………………………… and ……………………………………………………………………… is due to chance.

We <u>accept/reject</u> our null hypothesis.

Spearman's rank Question 3. Biodiversity on roundabouts



Roundabouts are common at road junctions in towns and cities. Ecologists investigated the species of plants and animals found on roundabouts in a small town. The grass on the roundabouts was mown at different time intervals. The table shows the mean number of plant species found on the roundabouts. From the data, is it possible to prove that mowing too frequently reduces biodiversity?

Approximate interval between mowing (days)	**Mean number of plant species**
7	15.8
14	8.3
40	21.2
80	30.6
365	32.0

Null hypothesis Write the null hypothesis here.

…………………………………………………………………………………………………..

…………………………………………………………………………………………………..

Rank the data & Calculate Spearman's rank

					$\sum D^2$ =

$\sum D^2$ = ...

n = ...

Substitute the numbers into the equation to calculate r_s

$$r_s = 1 - \left[\frac{6 \times \sum D^2}{n^3 - n}\right]$$

r_s = ...

<u>Interpret the results</u>

Our calculated value of Spearman's rank correlation coefficient is <u>smaller/bigger</u> than the critical value.

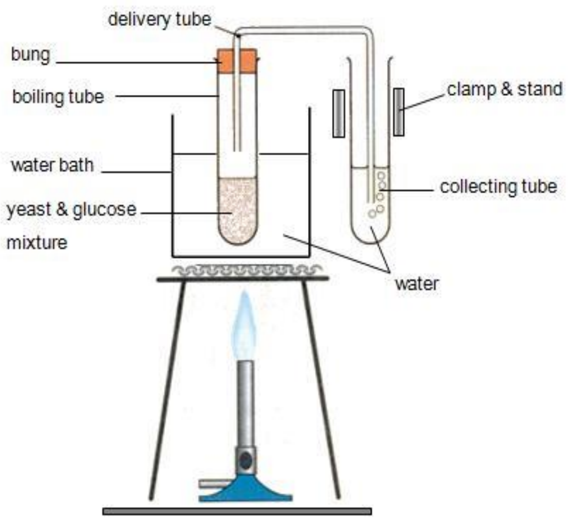
There is <u>less/more</u> than 5% probability that the <u>(positive/negative)</u> correlation between .... and .... is due to chance.

We <u>accept/reject</u> our null hypothesis.

More Spearman's rank correlation coefficient problems.

4. A student set up an experiment as shown opposite.

The water bath was heated to 30^{0}C, and the yeast left for 5 mins to allow the temperature of the yeast to equilibrate with the temperature of the water bath.
The rate of respiration in the yeast was then measured by recording the number of bubbles produced over 1 minute.
The experiment was then repeated at 40, 50, 60, 70, 80 and 90^{0}C.



The final results are shown below

Temperature (^{0}C)	Rate of respiration (Bubbles per minute)
10	2
20	5
30	9
40	17
50	32
60	12
70	4
80	0
90	0

What, if any, is the relationship between temperature and rate of respiration?
Use Spearman's Rank Correlation Coefficient (r_s) in your analysis of the results.

5. Great tits are small birds. In a study of growth in great tits, the relationship between the mass of the eggs and the mass of the young bird on hatching was investigated. Is there a relationship?

Egg mass / g	Chick mass / g
1.37	0.99
1.49	0.99
1.56	1.18
1.70	1.16
1.72	1.17
1.79	1.27
1.93	1.75

6. A student carried out an investigation to find out if there is a link between Mussel shell length and width on a rocky shore. Is there a relationship?

Shell length (mm)	Shell width (mm)
46	23
50	28
45	41
45	31
63	26
57	33
65	35
73	21
55	38
79	30
62	36
59	38
71	45
68	28
77	42

7. In a study of the 18 volunteers, the correlation between the mood and the amount of liquid consumed by daily drinking was investigated. The Spearman's rank correlation, $r_s = 0.12$ was obtained. How should this data be interpreted?

8. The correlation value obtained in a study of correlation between body height and biological age was $r_s = 0.97$. May we conclude that height and age are definitely excellently correlated?

9. A student carried out 6 samples examining the growth rate of bacteria and the concentration of citric acid present in the growth medium, and then calculated an r_s value of -0.89. How should this data be interpreted?

Student's t-test

Use this test when you are looking for the difference between two means, and you want to know if the difference is 'significant' or not. OK, so the formula looks a bit scary, but if you look through the worked examples you'll realise it's not so tough!

t can be calculated from the formula

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(s_1{}^2/n_1) + (s_2{}^2/n_2)}}$$

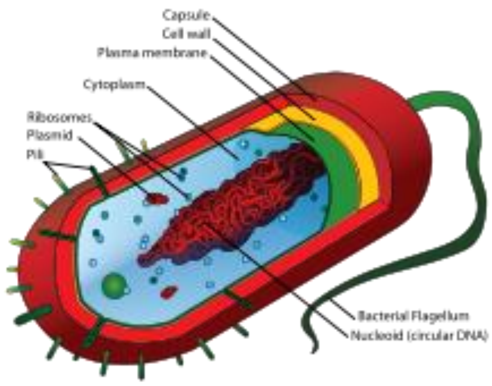
Where $\bar{x}_1$ = mean of first sample
$\bar{x}_2$ = mean of second sample
s_1 = standard deviation of first sample
s_2 = standard deviation of second sample
n_1 = number of measurements in first sample
n_2 = number of measurements in second sample

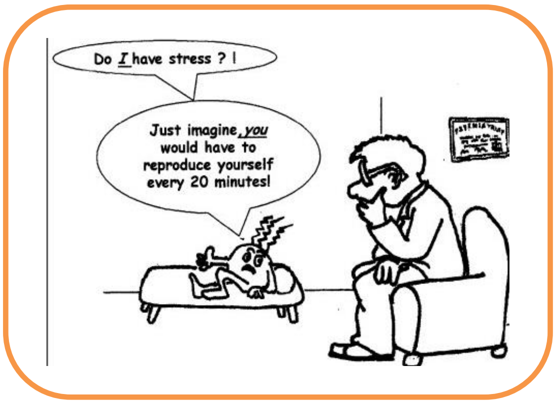
A table showing the critical values of *t* for different degrees of freedom.

Degrees of freedom	Critical value	Degrees of freedom	Critical value
4	2.78		
5	2.57	15	2.13
6	2.48	16	2.12
7	2.37	18	2.10
8	2.31	20	2.09
9	2.26	22	2.07
10	2.23	24	2.06
11	2.20	26	2.06
12	2.18	28	2.05
13	2.16	30	2.04
14	2.15	40	2.02

The number of degrees of freedom = $(n_1 + n_2) - 2$

Student's t-test   Worked Example 1. Bacteria





We have been growing two different strains of bacteria in flasks containing glucose. We had 4 replicate flasks for each bacterium. We have measured the biomass and want to find out whether or not the results are significantly different for the two different strains of bacteria.

	Mass (milligrams) of bacteria	
	Bacterium A	**Bacterium B**
Flask 1	520	230
Flask 2	460	270
Flask 3	500	250
Flask 4	470	280
Mean value	487.5	257.5

Null hypothesis

The null hypothesis when doing the Student t-test is

"there is no significant difference between the two different means"

Calculating the value of t

Next we calculate the value of t using the formula below.

t can be calculated from the formula

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(s_1{}^2/n_1) + (s_2{}^2/n_2)}}$$

Where $\bar{x}_1$ = mean of first sample

$\bar{x}_2$ = mean of second sample

s_1 = standard deviation of first sample

s_2 = standard deviation of second sample

n_1 = number of measurements in first sample

n_2 = number of measurements in second sample

The best way to show these calculations is in a table.

	Bacterium A	**Bacterium B**
Mean	487.5	257.5
N	4	4
s (standard deviation)	27.54	22.17
s^2	758.5	491.5
$s^2 \div n$	189.6	122.9

Substitute these values into the formula

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(s_1{}^2/n_1) + (s_2{}^2/n_2)}}$$

$$t = \frac{487.5 - 257.5}{\sqrt{(189.6 + 122.9)}}$$

$$t = \frac{230}{\sqrt{312.5}}$$

$$t = \frac{230}{\sqrt{17.7}}$$

$$t = 12.99$$

The ‘critical value’ & the ‘degrees of freedom’

Before we can interpret our results we need to work out the ‘critical value’. You will remember from Chi-squared that this represents the borderline between accepting or rejecting our null hypothesis. We get the critical value from the data sheet, but this depends on the number of ‘degrees of freedom’. Hopefully you will remember all about 'degrees of freedom' from Chi-squared. The calculation is slightly different simply because it allows for two sets of data.

A table showing the critical values of t for different degrees of freedom.

Degrees of freedom	Critical value	Degrees of freedom	Critical value
4	2.78		
5	2.57	15	2.13
6	2.48	16	2.12
7	2.37	18	2.10
8	2.31	20	2.09
9	2.26	22	2.07
10	2.23	24	2.06
11	2.20	26	2.06
12	2.18	28	2.05
13	2.16	30	2.04
14	2.15	40	2.02

The number of degrees of freedom = $(n_1 + n_2) - 2$

There were 4 flasks for each of the bacteria (n=4 for both bacteria)

Hence:

The number of degrees of freedom = (4 + 4) – 2

The number of degrees of freedom = 6

So we can see from the table of critical values of t, that 6 degrees of freedom = **2.48**

Our value of **t = 12.99** This is much higher than the critical value.

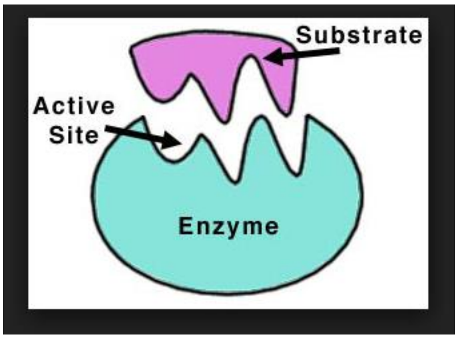
Interpreting the results

Our calculated value of t is **greater than** the critical value of t.

There is more than 5% probability that the differences in the means (mean mass of bacterium A and mean mass of bacterium B) are **not** due to chance.

We **reject** our null hypothesis.

Student's t-test   Worked Example 2. Enzymes



Some species of bacteria cause diseases of the stomach. Most are killed by acid gastric juices produced by the stomach lining. Some species of bacteria survive the antibacterial action of gastric juices by secreting the enzyme urease. This enzyme catalyses a reaction that produces ammonia. The ammonia neutralises the acid in gastric juice. A student believes that a small increase in temperature reduces the effect of urease and has produced the table of results below. The student wants to know whether her findings are significant or not.

Time for the acid to be neutralised (s)	
36.5°C	**37.5°C**
57	58
43	57
49	51
51	57
44	54
No data	49

Null hypothesis

The null hypothesis when doing the Student t-test is

"there is no significant difference between the two different means"

Calculating the value of t

	36.5°C	**37.5°C**
Mean	48.6	54.3
N	5	6
s (standard deviation)	5.68	3.67
s^2	32.26	13.47
$s^2 \div n$	6.5	2.2

Substitute these values into the formula

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(s_1{}^2/n_1) + (s_2{}^2/n_2)}}$$

$$t = \frac{48.6 - 54.3}{\sqrt{(6.5 + 2.2)}}$$

$$t = \frac{5.7}{\sqrt{8.7}}$$ (Ignore the minus sign, as the 'difference in means' is intended)

$$t = \frac{5.7}{2.95}$$

t = 1.9

<u>The 'critical value' & the 'degrees of freedom'</u>

A table showing the critical values of *t* for different degrees of freedom.

Degrees of freedom	Critical value	Degrees of freedom	Critical value
4	2.78		
5	2.57	15	2.13
6	2.48	16	2.12
7	2.37	18	2.10
8	2.31	20	2.09
9	2.26	22	2.07
10	2.23	24	2.06
11	2.20	26	2.06
12	2.18	28	2.05
13	2.16	30	2.04
14	2.15	40	2.02

The number of degrees of freedom = $(n_1 + n_2) - 2$

There were n = 5 for the lower temperature, but n = 6 for the higher temperature

<u>Hence:</u>

The number of degrees of freedom = (5 + 6) – 2

The number of degrees of freedom = 9

So we can see from the table of critical values of t, that 6 degrees of freedom = **2.26**

Our value of **t = 1.9** This is lower than the critical value.

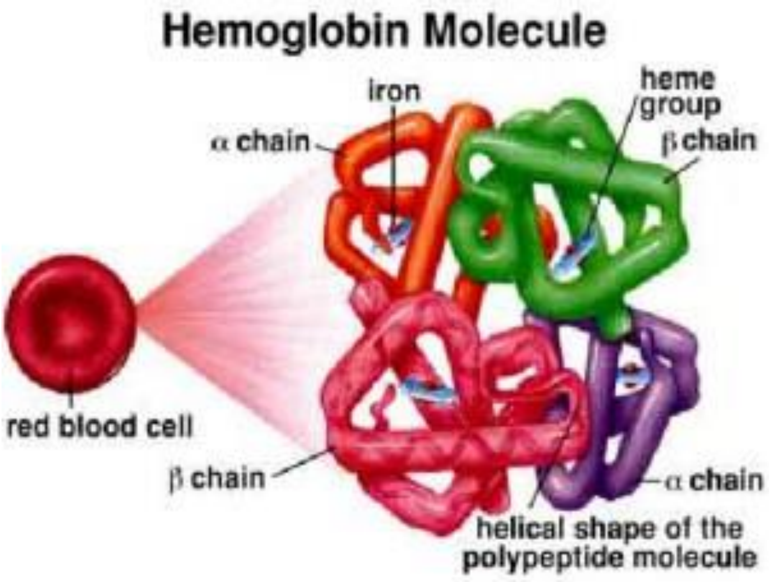
Interpreting the results

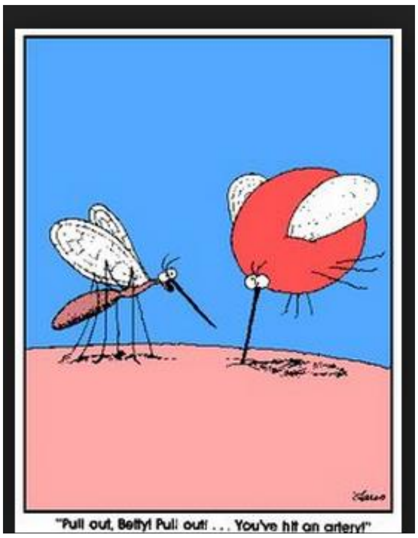
Our calculated value of t is **less than** the critical value of t.

There is more than 5% probability that the differences in the means (mean mass of bacterium A and mean mass of bacterium B) **are** due to chance.

We **accept** our null hypothesis.

Student's t-test Question 1.





A theatre nurse suspects that his newly purchased bit of expensive kit that gives instant blood - haemoglobin levels is faulty. So he takes readings and compares these with his trusty old bit of kit. His results are in the table below. Are his results significantly different?

Haemoglobin reading (g/dL)	
New expensive kit	**Old bit of kit**
17.6	16.5
13.9	12.5
15.4	14.8
16.5	15.2
13.5	12.1
16.5	15.3

The Null hypothesis is

………………………………………………………………………………………………………

………………………………………………………………………………………………………

Calculating the value of t

	New expensive kit	Old bit of kit
Mean		
N		
s (standard deviation)		
s^2		
$s^2 \div n$		

Substitute these values into the formula

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(s_1{}^2/n_1) + (s_2{}^2/n_2)}}$$

$$t = \frac{\quad - \quad}{\sqrt{(\quad + \quad)}}$$

$$t = \frac{\quad}{\sqrt{\quad}}$$

$$t = \frac{\quad}{\quad}$$

t =

The 'critical value' & the 'degrees of freedom'

Now calculate how many Degrees of freedom

The number of degrees of freedom = ……………………………

Use the table on the previous page to find the critical values of t = **……………………**

Interpreting the results

Our calculated value of t is **less/greater than** the critical value of t.

There is more than 5% probability that the differences in the means (mean mass of bacterium A and mean mass of bacterium B) **are/not** due to chance.

We **accept/reject** our null hypothesis.

Student's t-test   Question 2.

A scientist is examining the rate of mitosis in the root cells and shoot cells of a species of grass. She wants to know whether or not the rate of cells division in the root is quicker than the rate of mitosis in the shoot cells. Are her results significantly different?

Time for one cell cycle (hours)	
Shoot cells	**Root cells**
2.6	2.1
3.5	1.7
4.1	2.6
2.8	3.8
2.7	3.5
2.5	1.9
3.1	2.1
2.9	2.5
3.4	2.2

The Null hypothesis is

....

....

Calculating the value of t

	New expensive kit	Old bit of kit
Mean		
N		
s (standard deviation)		
s^2		
$s^2 \div n$		

Calculate the value of t

The number of degrees of freedom = ...

Use the table on the previous page to find the critical values of t = **.......................**

Interpreting the results

Our calculated value of t is **...**

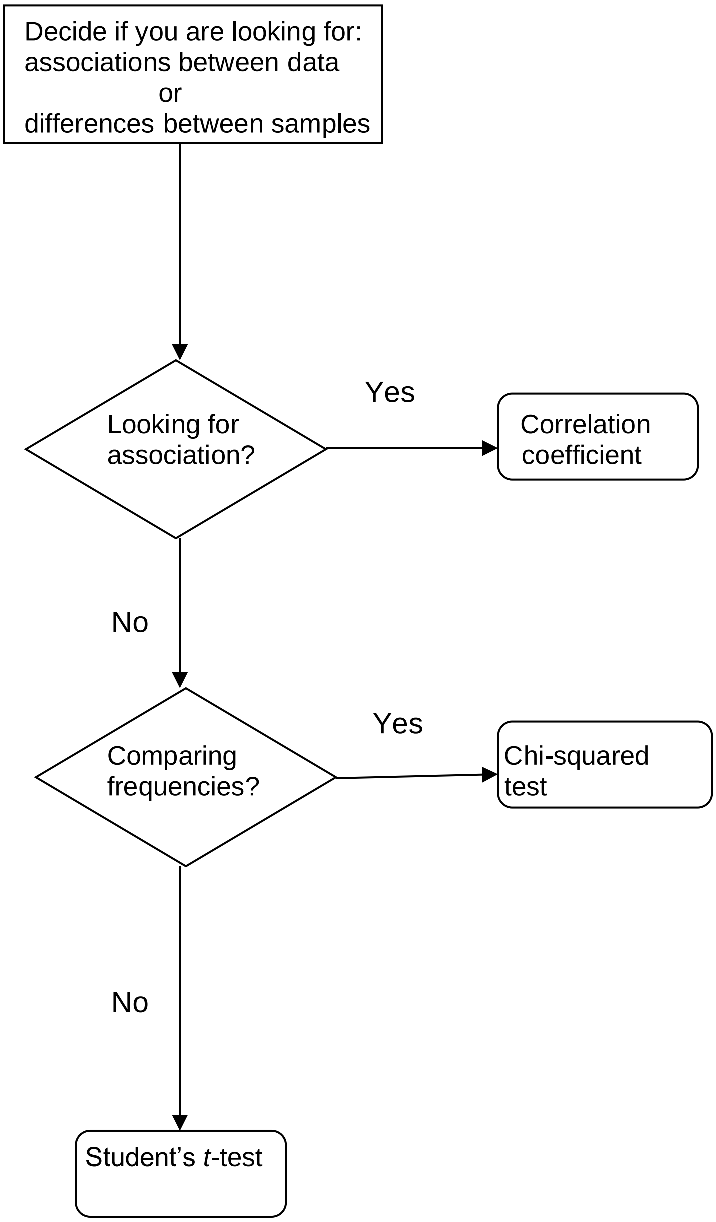
There is ....

We **..**

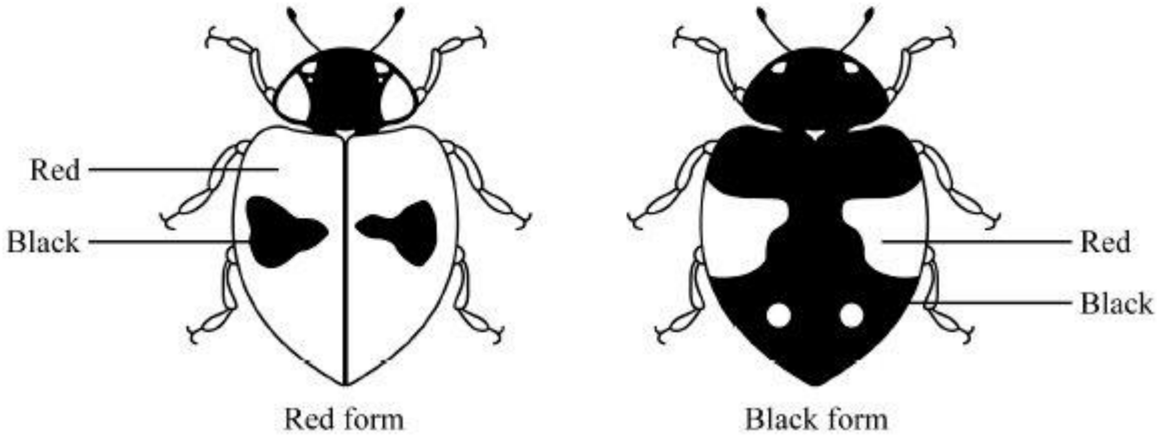
<u>Final Questions.</u>

For each of the following questions, use the flow diagram below to help you to use the most appropriate statistical test.

Flowchart for deciding which <u>Inferential</u> statistical test to use



1. The two-spot ladybird is a small beetle. It has a red form and a black form. These two forms are shown in the diagram.



Colour is controlled by a single gene with two alleles. The allele for black, B, is dominant to the allele for red, b. Scientists working in Germany compared the number of red and black ladybirds over a six-year period. They collected random samples of ladybirds from birch trees. Some of the results from the investigation are shown in the table.

How could you show that the frequency of the allele has remained the same?

Which statistical test should you use? Justify your choice of statistical test.

State the null hypothesis and interpret the results using the terms probability & chance.

Year	Season	Frequency of b allele
1933	Autumn	0.70
1934	Autumn	0.82
1935	Autumn	0.59
1936	Autumn	0.76
1937	Autumn	0.57
1938	Autumn	0.78

2. Fur seals live in Antarctic seas. They feed on fish and shrimp-like animals called krill. During the summer the fur seals come ashore to breed. The table shows the number of fur seals breeding on an Antarctic island from 1956 to 1986.

   How could the increase in adult fur seal numbers be shown to be significant?

   Which statistical test should you use? Justify your choice of statistical test.

   State the null hypothesis and interpret the results using the terms probability & chance.

Year	Number of adult fur seals
1956	100
1964	100
1970	200
1975	100
1976	1600
1981	2900
1983	3100
1986	11700

3. A young bird watcher was watching a pair of breeding blue tits bringing food back to the nest for the newly hatched chicks. He measured their 'return rate' which is a factor that takes into account the bird's time away from the nest, and the success in returning to the nest with food for the chicks. He created a table of his results. Is the male blue tit a significantly better provider for the chicks than the female?

Gender	Return rate (mg/hour)									
Female adult	56	75	45	71	61	64	58	80	76	61
Male adult	66	70	40	60	65	56	59	77	67	63

   Which statistical test should you use? Justify your choice of statistical test.

   State the null hypothesis and interpret the results using the terms probability & chance.

4. Malaria is a disease caused by a parasite. Scientists investigated the effect of malaria on competition between two species of Anolis lizard on a small Caribbean island. They sampled both populations by collecting lizards from a large number of sites on the island.

The scientists investigated the percentage of lizards of both species that were infected with malaria at different sites on the island. They collected samples of both lizards at intervals of 3 months for 1 year. They also recorded the elevation (height above sea level) of each site. Some of their results are shown in the table.

Site	**Elevation of collectio n site / metres**	**Total number of *A. gingivinus* collected in one year**	**Percentage of *A. gingivinus* infected with malaria**	**Total number of *A. wattsi* collected in one year**	**Percentage of *A. wattsi* infected with malaria**
1	10	13	0	0	0
2	80	30	0	0	0
3	120	35	23	3	0
4	200	40	30	7	0
5	300	52	46	12	0
6	315	35	31	13	1
7	370	155	37	79	2
8	414	124	44	68	4

(a) A preliminary study suggested that malarial infections in *A.gingivinus* were more common at higher elevations. Use the data provided to determine whether this suggestion is statistically significant.

Which statistical test should you use? Justify your choice of statistical test.

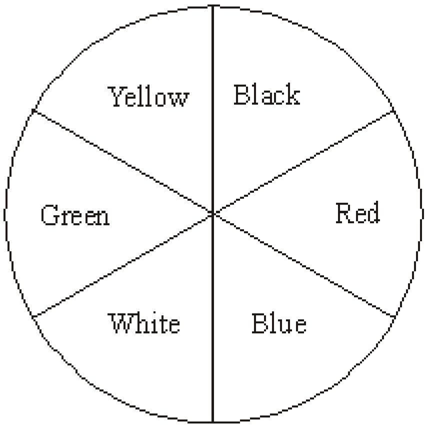
State the null hypothesis and interpret the results using the terms probability & chance.

(b) The scientists carried out a statistical test to determine whether the correlation between the number of A. wattsi collected and the percentage of A. gingivinus infected was significant.

Which statistical test should you use? Justify your choice of statistical test.

State the null hypothesis and interpret the results using the terms probability & chance.

5. In an investigation by a student into the responses of maggots, the bottom of a large box was marked with six coloured segments, as shown in the diagram.



30 maggots were placed on each segment in the box. A transparent cover was put on the box and light bulbs were positioned so that the segments were evenly illuminated. The positions of the maggots were recorded after one hour. The intensity of the light reflected by each segment was measured. The experiment was repeated three more times. The total number of maggots in each segment from the four experiments is shown in the table.

Colour of segment	Intensity of reflected light / arbitrary units	Total number of maggots
Black	4	154
Red	25	229
Blue	10	178
White	44	47
Green	25	48
Yellow	40	64

Give **one** conclusion about the responses of maggots which is supported by these results, and test your conclusion to see if it is statistically significant, using a suitable statistical test.

Which statistical test should you use? Justify your choice of statistical test.

State the null hypothesis and interpret the results using the terms probability & chance.

6. Here are the results of an investigation into the rate of photosynthesis in the pond weed Elodea. The number of bubbles given off in one minute was counted under different light intensities, and each measurement was repeated 5 times. How can you show that mean rate of photosynthesis at each light intensity is significantly different?

light intensity (Lux)	**Rate of photosynthesis (number of bubbles/min)**				
	repeat 1	**repeat 2**	**repeat 3**	**repeat 4**	**repeat 5**
0	5	2	0	2	1
500	12	4	5	8	7
1000	7	20	18	14	24
2000	42	25	31	14	38
3500	45	40	36	50	28
5000	65	54	72	58	36

7. In a test of two drugs 8 patients were given one drug and 8 patients another drug. The number of hours of relief from symptoms was measured with the following results:

Drug	Time spent symptom free (hours)							
A	3.2	1.6	5.7	2.8	5.5	1.2	6.1	2.9
B	3.8	1.0	8.4	3.6	5.0	3.5	7.3	4.8

Find out which drug is better by using an appropriate statistical test to find if it is significantly better than the other drug.

8. In one of Mendel's dihybrid crosses, the following types and numbers of pea plants were recorded in the F2 generation:

Number of Yellow round seeds	Number of Yellow wrinkled seeds	Number of Green round seeds	Number of Green wrinkled seeds
395	122	96	39

According to theory these should be in the ratio of 9:3:3:1.
Use the table of critical values below to determine whether these observed results agree with the expected ratio at P = 0.05 and P = 0.001?

Degrees of freedom	Probability of exceeding the critical value 0.10	0.05	0.025	0.01	0.001
1	2.706	3.841	5.024	6.635	10.828
2	4.605	5.991	7.378	9.210	13.816
3	6.251	7.815	9.348	11.345	16.266
4	7.779	9.488	11.143	13.277	18.467
5	9.236	11.070	12.833	15.086	20.515
6	10.645	12.592	14.449	16.812	22.458
7	12.017	14.067	16.013	18.475	24.322
8	13.362	15.507	17.535	20.090	26.125
9	14.684	16.919	19.023	21.666	27.877
10	15.987	18.307	20.483	23.209	29.588

9. The areas of moss growing on the north and south sides of a group of trees were compared.

Orientation	Total area of moss growing (m^2)					
North side of tree	20	43	53	86	70	54
South side of tree	63	11	21	54	9	74

Is there a significant difference between the north and south sides?

10. The table below shows the results of an experiment. Five different trays of seedlings were grown under red or yellow light over a four-hour period. The growth of the seedlings was measured. Are the differences in growth significant?

Tray of seedlings	**Mean increase in length/mm**	
	Grown in the red light	**Grown in yellow light**
P	5.2	3.8
Q	3.9	4.2
R	4.9	3.4
S	4.1	3.3
T	4.9	3.7