

## A Level Biology Statistics Summary

Test	Formula	Use	Degrees of Freedom	Accept/reject Null Hypothesis	Extra Information
Standard Deviation	$\sqrt{\frac{\sum(X - \bar{X})^2}{(n - 1)}}$ <p>where:  <math>X</math> = each score  <math>\bar{X}</math> = the mean or average  <math>n</math> = the number of values  <math>\Sigma</math> means we sum across the values</p>	Measure range of values around the mean	N/A	<p>There is <b>an/no</b> overlap in the +/- SD bars. This indicates the difference in the means of..... are <b>likely/unlikely</b> to be due to chance.</p> <p>Overlapping SD bars means the difference between two values is not significant.</p>	<p>The smaller the standard deviation, the narrower is the range, which translate to a higher reproducibility</p> <p>The small standard deviation means the experimental values are clustered together tightly (higher precision).</p>
T-Test	$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(s_1^2/n_1) + (s_2^2/n_2)}}$ <p>Where <math>\bar{x}_1</math> = mean of first sample  <math>\bar{x}_2</math> = mean of second sample  <math>s_1</math> = standard deviation of first sample  <math>s_2</math> = standard deviation of second sample  <math>n_1</math> = number of measurements in first sample  <math>n_2</math> = number of measurements in second sample</p>	To find the <b>difference between two means</b> and whether it is significant	$(n_1 + n_2) - 2$	If the t value is <b>less/greater</b> than the critical value, then there is 5% probability that the means ..... are/not due to chance. We <b>accept / reject</b> null hypothesis	<p>Less/due to chance /accept null hypothesis.</p> <p>Greater than critical value/ not due to chance/ reject null hypothesis</p>
Spearman's Rank correlation	$r_s = 1 - \frac{6 \sum D^2}{n(n^2 - 1)}$ <p><math>n^2</math>: Number of pairs  <math>D^2</math>: Difference between the results</p>	Used to see the strength of the correlation (association) <b>between two sets of continuous data</b>	n	If your calculated number is the same or higher than critical value then your correlation is significant and you would reject the null hypothesis	Remember to Rank the data The result will always be between 1 and minus 1.
Chi-Squared Test	$\chi^2 = \sum \frac{(O - E)^2}{E}$ <p>Where:  <math>\chi^2</math> = Chi Square obtained  <math>\Sigma</math> = the sum of  <math>O</math> = observed score  <math>E</math> = expected score</p>	Used to compare observed results with theoretical expected <b>frequencies</b> . The higher the $\chi^2$ value the greater the difference.	$(n-1)$	<p>If the calculated value is <b>greater/less</b> than critical value at 0.05 level then you can <b>accept/reject</b> the null hypothesis.</p> <p>Greater value than critical value/ reject null hypothesis</p>	Often used for genetics and fieldwork sampling.

### Notes

- Null Hypotheses: Always begins 'there is no significant difference/correlation between.....'
- Critical values: We generally use a value at the 0.05 or 95% probability (p) level. This means we accept there is a 5% probability that the results are due to chance
- If the calculated value exceeds the critical value then we can **reject** the null hypothesis

# Standard Deviation

Filey Brigg is a long, narrow peninsula in North Yorkshire. It has one shore exposed to the ocean and the other shore sheltered by a bay as shown in the photograph.



A limpet is an animal with a shell and is often found on these rocks. A limpet has a conical-shaped shell and a muscular foot that anchors it to the rock. The diagrams show possible shapes formed by a limpet on a rock.



(a) A study was carried out to investigate the effect of exposure on the shell height of limpets found on either side of Filey Brigg.

The table shows the data collected from the exposed shore.

Limpet	Shell Height / mm	$(x - \bar{x})$	$(x - \bar{x})^2$
1	9.0	0.7	0.45
2	5.9	-2.4	5.90
3	8.3	0.0	0.00
4	11.6	3.3	10.69
5	8.4	0.1	0.00
6	7.8	-0.5	0.28
7	9.2	0.9	0.76
8	9.1	0.8	0.59
9	10.0	1.7	2.79
10	7.9	-0.4	0.18
11	5.0	-3.3	11.09
12	9.3	1.0	0.94
13	9.1	0.8	0.59
14	7.8	-0.5	0.28
15	7.3	-1.0	1.06
16	8.6	0.3	0.07
17	6.9	-1.4	2.04
18	8.4	0.1	0.00
19	7.9	-0.4	0.18
20	9.1		
	$\bar{x} = 8.33$		



(i) Give a biotic factor that could affect the height of limpet shells.

(1)

(ii) Calculate the standard deviation of the exposed shore using the formula provided.

(2)

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

Answer.....

13

Turn over ▶



S 5 8 7 0 1 A 0 1 3 3 2

DO NOT WRITE IN THIS AREA

(b) The same number of limpets were measured on the sheltered shore. The null hypothesis for this study was that there will be no difference between the height of the limpet shells from the sheltered and exposed shores of Filey Brigg. A paired t-test was performed on the collected data. The t value was calculated to be 2.095.

The table shows some of the critical values for the paired t-test.

Degrees of freedom	p = 0.1	p = 0.05	p = 0.01
18	1.734	2.101	2.878
19	1.729	2.093	2.861
20	1.725	2.086	2.845
21	1.721	2.080	2.831

A student concluded from these data that the null hypothesis should be accepted. Comment on the validity of this conclusion.

(3)

(Total for Question 5 = 6 marks)

14



S 5 8 7 0 1 A 0 1 4 3 2

DO NOT WRITE IN THIS AREA

44 A student investigated the effect of salt concentration on the growth of one species of brine shrimp.

The student placed 100 shrimp eggs in a beaker containing 1 dm<sup>3</sup> of 3% salt solution. Three days after the eggs hatched, 10 shrimps were collected and their lengths measured. Seven days after hatching, another 10 shrimps were collected and their lengths measured.

The procedure was repeated using a 5% salt solution. All other variables were kept constant. The results are shown in the table.

Specimen number	Length of specimen in 3% salt solution / mm		Length of specimen in 5% salt solution / mm	
	3 days after hatching	7 days after hatching	3 days after hatching	7 days after hatching
1	0.75	1.00	0.75	0.98
2	0.78	1.25	0.73	0.95
3	0.66	1.10	0.61	0.93
4	0.73	1.03	0.63	0.83
5	0.85	1.15	0.53	0.98
6	0.78	1.08	0.60	1.08
7	0.90	1.13	0.52	0.95
8	0.90	1.05	0.81	1.03
9	0.80	1.18	0.58	0.88
10	0.85	1.05	0.68	0.88
Mean	0.800	1.102		0.949

- (a) (i) Calculate the mean length of shrimp three days after hatching in the 5% salt solution. Give your answer to an appropriate number of significant figures. (1)

Answer ..... mm

DO NOT WRITE IN THIS AREA

- (ii) Explain how these data can be used to show the effect of salt concentration on the rate of growth of brine shrimps. (2)

DO NOT WRITE IN THIS AREA

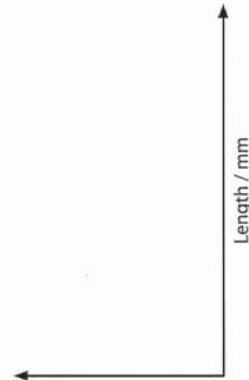
- (b) The student carried out a statistical test to see if there was a significant difference between the mean lengths of brine shrimps in these two salt solutions. The student used the 7 days after hatching data.

The student selected the t-test because the data are normally distributed.

- (i) Draw a line on the graph to show a normal distribution for a population of brine shrimp. (1)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA



(ii) Complete the table by filling in the missing value for specimen 7 and then calculate the value for  $\sum x_1^2$ .

(1)

Specimen number (n)	3% salt solution		5% salt solution	
	Length of specimen ( $x_1$ )	Square of length of specimen ( $x_1^2$ )	Length of specimen ( $x_2$ )	Square of length of specimen ( $x_2^2$ )
1	1.00	1.000	0.98	0.960
2	1.25	1.563	0.95	0.903
3	1.10	1.210	0.93	0.865
4	1.03	1.061	0.83	0.689
5	1.15	1.323	0.98	0.960
6	1.08	1.166	1.08	1.166
7	1.13	.....	0.95	0.903
8	1.05	1.103	1.03	1.061
9	1.18	1.392	0.88	0.774
10	1.05	1.103	0.88	0.774
Sum ( $\Sigma$ )	$\Sigma x_1 = 11.020$	$\Sigma x_1^2 =$ .....	$\Sigma x_2 = 9.490$	$\Sigma x_2^2 = 9.055$
Mean	$\bar{x}_1 = 1.102$		$\bar{x}_2 = 0.949$	

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

(iii) The variances are used in the calculation of a t-value.

The variance for the 3% salt solution  $S_1^2 = 0.0059$ .

Calculate the variance for the 5% salt solution ( $S_2^2$ ) using the formula

$$S_2^2 = \frac{\sum x_2^2 - \frac{(\sum x_2)^2}{n}}{n - 1}$$

n = the number of specimens

(2)

$S_2^2$  .....

(iv) Calculate the t-value using the formula.

Give your answer to an appropriate number of significant figures.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

(3)

t = .....

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA



Speeman Rank

DCMU is a herbicide that can disrupt one of the carrier proteins in the electron transport chain of chloroplasts.

A student carried out an investigation to study the effect of DCMU concentration on the Hill reaction.

The student ground up some spinach leaves in an isolation mixture containing sucrose solution at a concentration of  $0.4 \text{ mol dm}^{-3}$ . The mixture was filtered and then spun in a centrifuge. The chloroplasts were extracted.

These chloroplasts were divided equally into eight different tubes containing a solution of distilled water, DCPIP and a buffer.

A small volume of DCMU was added to each tube and the time taken for the blue DCPIP to decolourise was recorded.

- (a) (i) Explain why the isolation mixture contained sucrose at a concentration of  $0.4 \text{ mol dm}^{-3}$  rather than at a concentration of  $0.8 \text{ mol dm}^{-3}$ . (2)

.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....

- (ii) The isolation mixture had a temperature of  $4^\circ\text{C}$  and pH of 7. Explain why this temperature and pH are appropriate. (3)

.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....

- (b) The time taken for the DCPIP to decolourise was converted to a rate of colour change. The results are shown in the table.

Concentration of DCMU / $\mu\text{mol dm}^{-3}$	Rate of colour change / $\times 10^{-3} \text{ s}^{-1}$
0	23.5
1	12.5
2	6.3
3	4.2
4	3.2
5	2.6
6	2.2
7	2.3

- (i) Analyse the data to deduce a null hypothesis for this investigation. (1)

.....  
.....  
.....

DO NOT WRITE IN THIS AREA



- (ii) The student started to analyse the data using a statistical test, called the Spearman's rank correlation coefficient ( $r_s$ ), to see if there was a correlation.

Complete the table by filling in the empty boxes.

(1)

Concentration of DCMU / $\text{mol dm}^{-3}$	Concentration ranked ( $R_1$ )	Rate of colour change / $10^{-5} \text{ s}^{-1}$	Rate ranked ( $R_2$ )	Difference in ranks d ( $R_1 - R_2$ )	Difference squared $d^2$
0	1	23.5	8		
1	2	12.5	7	5	25
2	3	6.3	6	3	9
3	4	4.2	5	1	1
4	5	3.2	4	1	1
5	6	2.6	3	3	9
6	7	2.2	1	6	36
7	8	2.3	2	6	36

- (iii) Calculate  $r_s$  by using the equation.

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

Where  $\sum(d^2)$  is the sum of all the values for  $d^2$  and  $n$  is the number of pairs of data.

(3)

$r_s =$  .....

- (iv) The table shows some of the critical values for the Spearman's rank correlation coefficient.

Number of pairs of data ( $n$ )	p value			
	0.10	0.05	0.02	0.01
6	0.829	0.866	0.943	1.000
8	0.643	0.786	0.833	0.881
10	0.564	0.648	0.746	0.794
12	0.506	0.591	0.712	0.777
14	0.456	0.544	0.645	0.715
16	0.425	0.506	0.601	0.665
18	0.399	0.475	0.564	0.625

Explain the outcome of the statistical test.

(3)

(Total for Question 3 = 13 marks)

DO NOT WRITE IN THIS AREA



2 Blowfly larvae can be used by a forensic scientist to help determine the time of death of a body.

The diagram shows a Petri dish used by a student to investigate whether young and old blowfly larvae show a preference for light or dark conditions.



In the first trial, the left side was dark and the right side was light.

Five blowfly larvae were added to each side of the chamber.

After five minutes, the number of larvae on each side of the Petri dish was recorded.

In the second trial, the same experiment was repeated but this time the right side was dark and the left side was light.

The table shows the results of the trials.

Trial	Number of young blowfly larvae		Number of old blowfly larvae	
	Left side dark	Right side light	Left side dark	Right side light
1	9	1	2	8
2	2	8	9	1

(a) Give a null hypothesis for this investigation.

(1)

(b) The Chi squared test can be used to determine whether the results of this investigation indicate a significant difference in the distribution of young larvae between the light and the dark side.

(i) Use the formula to calculate the Chi-squared value for young larvae.

$$\chi^2 = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}} \quad (3)$$

Answer .....

(ii) The table below gives some critical values for Chi-squared.

p value		
0.15	0.1	0.05
2.07	2.71	3.84
		5.02

Use your calculated value to determine whether the difference between the observed and expected results is significant.

(1)



Spearman Rank

Question Number	Answer	Additional Guidance	Mark
4(b)(iii)	<ul style="list-style-type: none"> <li>Top line of formula correctly calculated (1)</li> <li>correct answer to two significance figures (1)</li> </ul>	Example of calculation $9.055 - 9.006 = 0.0054 / 0.00544 / 5.4 \times 10^{-3}$	(2)
Question Number 4(b)(iv)	<ul style="list-style-type: none"> <li>top line of formula correctly calculated (1)</li> <li>bottom line of formula correctly calculated (1)</li> <li>correct answer to between three and five significant figures (1)</li> </ul>	Additional Guidance Example of calculation 0.153 0.0336 (OR 0.0337 if 0.00544 used) = 4.55 (OR 4.54 if 0.00544 used) ALLOW answer between 4.5510 and 4.5540 Correct answer with no working gains full marks	(3)

Question Number	Answer	Additional Guidance	Mark
4(b)(v)	An explanation that makes reference to the following: <ul style="list-style-type: none"> <li>there was a significant difference between {the 3% and the 5% salt solution / groups} (1)</li> <li>at the 5% significance level (1)</li> </ul>	IGNORE significant correlation / significant relationship ALLOW 95% probability there is a difference e.g. '5% chance that the difference is due to chance' or 'with 95% certainty' IGNORE $p = 0.05$	(2)

Question Number	Acceptable Answer	Additional Guidance	Mark
3(a)(i)	An explanation that makes reference to the following: <ul style="list-style-type: none"> <li>there is no concentration gradient present between the chloroplast and the isolation solution (1)</li> <li>no net loss of water from the chloroplast from osmosis (1)</li> </ul>		(2)
Question Number 3(a)(ii)	Acceptable Answer An explanation that makes reference to the following: <ul style="list-style-type: none"> <li>low temperature to (temporarily) slow enzyme activity (1)</li> <li>pH 7.0 so enzymes not denatured (1)</li> <li>so RUBISCO remains active (1)</li> </ul>	Additional Guidance	(3)
Question Number 3(b)(i)	Acceptable Answer	Additional Guidance	Mark
3(b)(i)	There is no correlation between the concentration of DCMU and the rate of DCPIP colour change		(1)
Question Number 3(b)(ii)	Acceptable Answer	Additional Guidance	Mark
3(b)(ii)	7 and 49		(1)

Question Number	Acceptable Answer	Additional Guidance	Mark
3(b)(iii)	<ul style="list-style-type: none"> <li>Correct calculation of numerator (1)</li> <li>Correct calculation of denominator (1)</li> <li>Correct calculation of correlation coefficient (1)</li> </ul>	Example of calculation: $(\sum d^2 = 996) \div (n(n^2-1) = 504) = (-) 0.976$ Allow all marks for correct answer with no working	(3)
Question Number 3(b)(iv)	Acceptable Answer	Additional Guidance	Mark
3(b)(iv)	An explanation that makes reference to the following: <ul style="list-style-type: none"> <li>Selection of appropriate critical value from the table (1)</li> <li>calculated value is greater (than critical value) (1)</li> <li>Can reject the null hypothesis / correlation is significant (1)</li> </ul>	0.786, 0.833, 0.881	(3)

Total for Question 3 = 13 marks

### Chi-Squared ( $\chi^2$ )

Question Number	Acceptable Answer	Additional guidance	Mark
2(a)	larvae show no significant preference for light over dark side (1)	Allow vice versa Must have NO in hypothesis.	(1)
2(b)(i)	Acceptable Answer calculation of expected frequency 10 and 10 (1) $(O-E)^2/E$ for both light and dark sides $49 + 10 = 4.9$ (1) sum = 9.8 (1)	Additional guidance	Mark (3)
2(b)(ii)	Acceptable Answer An answer that makes reference to the following: <ul style="list-style-type: none"> <li>higher than 3.84 therefore Chi square value as high as 9.8 arise by chance alone less than 1 in 20 / 0.05 therefore there is a significant difference (1)</li> </ul>	Additional guidance allow ECF for incorrect value of Chi allow converse if calculated of Chi is lower than 3.84	Mark (1)
2(c)	Acceptable Answer An explanation that makes reference to five of the following: <ul style="list-style-type: none"> <li>use of dead tissue containing toxin and dead tissue not containing tissue so valid comparison can be made (1)</li> <li>{young larvae of same length / fly eggs} allowed access to both types of dead tissue so they have same potential for growth (1)</li> <li>reference to time scale before growth of larvae measured to allow time for growth to occur (1)</li> <li>length measured for several larvae to ensure reliability (1)</li> <li>control of {temperature / type of tissue / age of tissue / species of larvae} because these factors affect growth (1)</li> <li>recognition that comparison of results may show under or over estimate of time of death (1)</li> </ul>	Additional guidance	Mark (5)

### Hardy Weinberg

Question Number	Acceptable Answer	Additional Guidance	Mark
5(a)(i)	location (of gene) on a chromosome		(1)
5(a)(ii)	Acceptable Answer An answer that makes reference to the following: <ul style="list-style-type: none"> <li>greater mean genetic diversity in East Thailand (1)</li> <li>the range of data shows there is overlap in the number of different alleles per gene locus (1)</li> <li>lowest (mean number of different alleles) for East Thailand is 1.30 and highest for South Thailand is 1.32 (1)</li> <li>unlikely to be a significant difference between the two regions (1)</li> </ul>	Additional Guidance	Mark (4)
5(b)(i)	Acceptable Answer An answer that makes reference to the following: <ul style="list-style-type: none"> <li>homozygous for allele 1 (1)</li> <li>only allele 1 is present /100% of the alleles are allele 1 (1)</li> </ul>	Additional Guidance	Mark (2)
5(b)(ii)	Acceptable Answer correct proportion of heterozygotes (2pq) (1) correct number of heterozygotes (1)	Example of calculation: $2 \times (0.980 \times 0.02) = 0.0392$ $0.0392 \times 50 = 1.96$ $= 2$	Mark (2)
5(b)(iii)	Acceptable Answer An answer that makes reference to three of the following: <ul style="list-style-type: none"> <li>allele 3 is result of a mutation (1)</li> <li>(allele 3) confers a selective advantage (1)</li> <li>population Z isolated from other populations (1)</li> <li>it is rare therefore it will not always be present in all populations (1)</li> </ul>	Additional Guidance	Mark (3)