Standard Error and 95% Confidence Limits

Example There is no significant difference between the mean shell lengths of mussels from the two locations.

	Area 1	Area 2
$\bar{x}$	57.8	30.6
SD	12.0	6.38
SE	3.79	2.02
CL lower	50.2	26.6
CL upper	65.4	34.6

The 95% confidence limits do not overlap, therefore we reject the null hypothesis. The probability that the differences in mean results are due to chance is less than or equal to 5%.

1. There is no significant difference between the changes in pH for different types of milk.

	Full fat	Semiskimmed	UHT
$\bar{x}$	1.29	0.813	0.900
SD	0.337	0.364	0.320
SE	0.112	0.129	0.101
CL lower	1.07	0.555	0.698
CL upper	1.51	1.07	1.10

The 95% confidence limits overlap, therefore we accept the null hypothesis. The probability that the differences in mean results are due to chance is greater than 5%.

2. There is no significant difference between the rates of flow of a river on different days.

	Day 1	Day 2
$\bar{x}$	265	249
SD	17.5	24.8
SE	5.53	7.84
CL lower	254	233
CL upper	276	265

The 95% confidence limits overlap, therefore we accept the null hypothesis. There is no significant difference between the rates of flow in the river on different days.

3. There is no significant difference between the numbers of beetles collected in two woodlands.

	Deciduous	Coniferous
$\bar{x}$	7.50	5.40
SD	2.42	2.37
SE	0.764	0.748
CL lower	5.97	3.90
CL upper	9.03	6.90

The 95% confidence limits overlap, therefore we accept the null hypothesis. There is no significant difference between the numbers of beetles collected in deciduous and coniferous woodland

Spearman's Rank

Example There is no correlation between the mass of nitrogen applied to fields and the concentration of nitrates in the streams.

Mass of N ($kg\,ha^{-1}$)	Rank	Conc. of nitrates ($mg\,dm^{-3}$)	Rank	D	D^2
41	1.5	1.2	1	0.5	0.25
41	1.5	1.3	2	-0.5	0.25
51	3	1.5	3	0	0
56	4	1.8	5	1	1
63	5	1.6	4	1	1
69	6	1.9	6	0	0
72	7	2.0	7	0	0

$\Sigma D^2 = 2.50$

$$r_s = 1 - \frac{6 \times 2.50}{7^3 - 7} = 1 - 0.0446 = \underline{\underline{0.955}}$$

The calculated value for r_s is greater than the critical value for 7 pairs of measurements (0.79), therefore we reject the null hypothesis. There is less than or equal to a 5% probability that the correlation in results occurred by chance, there is significant positive correlation between the mass of nitrogen applied to fields and the concentration of nitrates in nearby streams

1. There is no correlation between the concentration of enzyme and the rate of reaction

Enzyme conc. (mM)	Rank	Rate of reaction (a.u.)	Rank	D	D^2
0	1	0	1	0	0
0.1	2	0.8	2	0	0
0.2	3	1.1	3	0	0
0.4	4	1.8	4	0	0
0.5	5	3.2	5	0	0
0.8	6	6.6	6	0	0
1.0	7	7.2	7	0	0

$\Sigma D^2 = 0$

$$r_s = 1 - \frac{6 \times 0}{7^3 - 7} = 1 - 0 = \underline{\underline{1.00}}$$

The value for r_s is greater than the critical value for 7 pairs of measurements (0.79), therefore we reject the null hypothesis. There is less than or equal to a 5% probability that the correlation in results occurred by chance; there is significa positive correlation between the enzyme concentration and the rate of reaction.

Spearman's Rank

2 There is no correlation between the widths and heights of limpets.

Width (mm)	Rank	Height (mm)	Rank	D	D^2
22	5.5	29	2	3.5	12.25
23	7	39	8	-1	1
26	8	34	4	4	16
20	4	34	4	0	0
11	1	34	4	-3	9
18	3	40	9	-6	36
27	9	37	6.5	2.5	6.25
22	5.5	37	6.5	-1	1
29	10	17	1	9	81
15	2	47	10	-8	64

$\Sigma D^2 = 226.5$

$$r_s = 1 - \frac{6 \times 226.5}{10^3 - 10} = 1 - 1.373 = \underline{\underline{-0.373}}$$

The value of r_s is lower than the critical value for 10 pairs of measurements (0.65), therefore we accept the null hypothesis. There is greater than a 5% probability that any correlation in the results has occurred by chance; there is no significant correlation between the width and height of limpets.

3. There is no correlation between the percentage cover of bilberry and that of common heather

% cover bilberry	Rank	% cover heather	Rank	D	D^2
5	6.5	0	4	2.5	6.25
40	11	0	4	7	49
50	12	5	11	1	1
5	6.5	0	4	2.5	6.25
10	8.5	0	4	4.5	20.25
25	10	0	4	6	36
0	2	1	9.5	-7.5	56.25
4	5	0	4	1	1
0	2	0	4	-2	4
0	2	1	9.5	-7.5	56.25
10	8.5	6	12	-3.5	12.25
2	4	0.5	8	-4	16

$\Sigma D^2 = 264.5$

$$r_s = 1 - \frac{6 \times 264.5}{12^3 - 12} = 1 - 0.925 = \underline{\underline{0.0752}}$$

The value of r_s is lower than the critical value for 12 pairs of measurements (0.59) therefore we accept the null hypothesis. There is greater than a 5% probability that the correlation in results has occurred by chance; there is no significant correlation between the percentage cover of bilberry and the percentage cover of common heather.

X^2

Example There is no difference between the death rates on the North and South sides of the river.

Side of river	O	E	O-E	$(O-E)^2$	$\frac{(O-E)^2}{E}$
North	26	19	7	49	2.58
South	12	19	-7	49	2.58

$$X^2 = \sum \frac{(O-E)^2}{E} = 2.58 + 2.58 = \underline{5.16}$$

The value for X^2 is greater than the critical value for 1 degree of freedom (3.84), therefore we reject the null hypothesis. There is less than or equal to a 5% probability that the differences in results occurred by chance; there is a significant difference between the death rate on the North and South sides of the river.

1. There is no difference between the number of woodlice in different environments.

Environment	O	E	O-E	$(O-E)^2$	$\frac{(O-E)^2}{E}$
Dry, dark	22	25	-3	9	0.36
Dry, light	26	25	1	1	0.04
Wet, dark	31	25	6	36	1.44
Wet, light	21	25	-4	16	0.64

$$X^2 = \sum \frac{(O-E)^2}{E} = 0.36 + 0.04 + 1.44 + 0.64 = \underline{2.48}$$

The value for X^2 is less than the critical value for 3 degrees of freedom (7.82), therefore we accept the null hypothesis. There is greater than a 5% probability that the differences in the results occurred by chance; there is no significant difference between the number of woodlice found in each environment.

2. There is no significant difference between the numbers of moths of each colour caught.

Moth colour	O	E	O-E	$(O-E)^2$	$\frac{(O-E)^2}{E}$
Light	17	31	-14	196	6.32
Dark	45	31	14	196	6.32

$$X^2 = \Sigma \frac{(O-E)^2}{E} = 6.32 + 6.32 = \underline{\underline{12.64}}$$

The value for X^2 is greater than the critical value for 1 degree of freedom (3.84), therefore we reject the null hypothesis. There is less than or equal to a 5% probability that the differences in results occurred by chance; there is a significant difference between the numbers of light and dark moths caught in a polluted woodland.

3. There is no significant difference between the number of peas of each phenotype and their expected ratios.

Phenotype	O	E	O-E	$(O-E)^2$	$\frac{(O-E)^2}{E}$
Yellow, round	289	307.125	-18.125	328.51...	1.070
Yellow, wrinkled	122	102.375	19.625	385.14...	3.762
Green, round	96	102.375	-6.375	40.64...	0.3970
Green, wrinkled	39	34.125	4.875	23.76	0.6964

$$X^2 = \Sigma \frac{(O-E)^2}{E} = 1.070 + 3.762 + 0.3970 + 0.6964 = 5.93$$

$\underline{X^2}$

3. The value for X^2 is less than the critical value for 3 degrees of freedom (7.82) therefore we accept the null hypothesis. There is greater than a 5% probability that the differences between the observed and expected results occured by chance; there is no significant difference between the number of peas of each phenotype and the expected number of peas of each phenotype.

Analysing results using statistical methods - Revision

1\. a) There is no significant difference between the heights of seedlings when treated with different fertilisers.
b) Standard error and 95% confidence limits
c) We are looking for differences between mean values.
d)

	A	B	C
$\bar{x}$	0.594	0.678	0.448
SD	0.117	0.138	0.140
SE	0.0523	0.0615	0.0625
CL lower	0.489	0.555	0.323
CL upper	0.699	0.801	0.573

e) The 95% confidence limits for all three sets of data overlap, therefore we accept the null hypothesis. There is greater than a 5% probability that the differences in results occurred due to chance; there is no significant difference between the heights of seedlings grown with different fertilisers.

2\. a) There is no correlation between the conductivity and the diversity of a river.
b) Spearman's Rank
c) We are looking for associations between different measurements from the same sample.
d)

Conductivity (μS)	Rank	Diversity index	Rank	D	D^2
413.3	1	6.01	7	-6	36
588.5	4	6.50	8	-4	16
706.6	5	4.23	3	2	4
439.7	2	5.17	5	-3	9
726.0	6	4.49	4	2	4
850.0	8	3.82	2	6	36
567.3	3	5.68	6	-3	9
766.7	7	3.74	1	6	36

$r_s = 1 - \frac{6 \times \sum D^2}{n^3 - n} = 1 - \frac{6 \times 150}{8^3 - 8} = 1 - 1.786 = \underline{-0.786}$

e) The value of r_s is greater than the critical value for 8 pairs of measurements (0.74), therefore we reject the null hypothesis. There is less than or equal to a 5% probability that the correlation in results occurred by chance; there is a significant negative correlation between the conductivity and the diversity index of the river.

3.a) There is no significant difference between the numbers of birds feeding with each feeder.

b) χ^2

c) The investigation involves frequencies

d)

Feeder attachment	O	E	O−E	$(O-E)^2$	$\frac{(O-E)^2}{E}$
Elastic bands	53	42	11	121	2.88
String	43	42	1	1	0.0238
Fence post	30	42	−12	144	3.43

$\chi^2 = \sum \frac{(O-E)^2}{E} = 2.88 + 0.0238 + 3.43 = \underline{6.33}$

e) The value of χ^2 is greater than the critical value for 2 degrees of freedom (5.99), therefore we reject the null hypothesis. There is a less than or equal to 5% probability that the differences in results occurred by chance; there is a significant difference between the numbers of birds using each type of feeder.

4. a) There is no significant difference between the numbers of flat periwinkles found on different species of seaweed

b) χ^2

4. c) The investigation involves frequencies.

d)

Seaweed species	O	E	O-E	$(O-E)^2$	$\frac{(O-E)^2}{E}$
F. serratus	45	20	25	625	31.25
F. vesiculosus	38	20	18	324	16.2
A. nodosum	10	20	-10	100	5
F. spiralis	5	20	-15	225	11.25
Other	2	20	-18	324	16.2

$$X^2 = \sum \frac{(O-E)^2}{E} = 31.25 + 16.2 + 5 + 11.25 + 16.2 = 79.9$$

e) The value of X^2 is greater than the critical value for 4 degrees of freedom (9.49), therefore we reject the null hypothesis. There is less than or equal to a 5% probability that the differences in results are due to chance; there is a significant difference between the numbers of flat periwinkle found on different species of seaweed.

5. a) There is no difference between the areas of moss growing on the north and south sides of trees.

b) Standard error and 95% confidence limits

c) Looking for differences between mean values.

d)

	North	South
$\bar{x}$	52.3	37.5
SD	18.0	21.5
SE	5.70	6.80
CL lower	40.9	23.9
CL upper	63.7	51.1

e) The 95% confidence limits overlap, therefore we accept the null hypothesis. There is greater than a 5% probability that the differences in results occurred by chance; there is no significant difference between the areas of moss growing on the north and south sides of trees.

6.a) There is no correlation between the phosphate and chlorophyll concentrations of the lakes.

b) Spearman's rank.

c) Looking for associations between different measurements from the same sample

d)

Phosphate conc (gdm^{-3})	Rank	Chlorophyll conc. (gdm^{-3})	Rank	D	D^2
340	7	460	7	0	0
200	4	230	3	1	1
320	6	410	6	0	0
240	5	370	4.5	0.5	0.25
160	1	220	2	-1	1
190	2.5	370	4.5	-2	4
190	2.5	200	1	1.5	2.25

$$r_s = 1 - \frac{6 \times \Sigma D^2}{n^3 - n} = 1 - \frac{6 \times 8.5}{7^3 - 7} = 1 - 0.152 = \underline{0.848}$$

e) The value of r_s is greater than the critical value for 7 pairs of measurements (0.79), therefore we reject the null hypothesis. There is less than or equal to a 5% probability that the correlation in the results occurred due to chance; there is significant positive correlation between the phosphate and chlorophyll concentrations in the lakes.