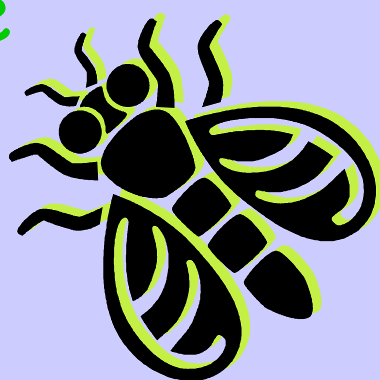
THE CHI-SQUARED TEST

Probability and Chance

Breeding experiments like Mendel's give the expected ratios as long as:-

- Fertilisation is random
- There is an equal opportunity of survival among the offspring
- Large numbers of offspring are produced



If these criteria are not met the observed results of a breeding experiment may not necessarily agree with the initial prediction. For example,

- there is a chance that more pollen grains of one genetic constitution may fuse with the egg cells than another
- if the cross produces very few progeny, the progeny may not appear in the expected proportions
- the chance of a 3:1 ratio occurring as a result of a monohybrid cross, increases as more progeny are produced

Chi-squared test

- If the actual test result is not the same as the prediction, is the difference due to chance, or is the prediction wrong?
- The larger the difference between the observed and predicted outcomes, the more likely it is that the prediction is wrong

- The chi-squared test is used to estimate the probability that the difference between the observed and expected results are due to chance.
- It does this by measuring the size of the difference (deviation) between the observed results (O) and the expected results (E)
- It can only be used to test discontinuous (discrete) variables

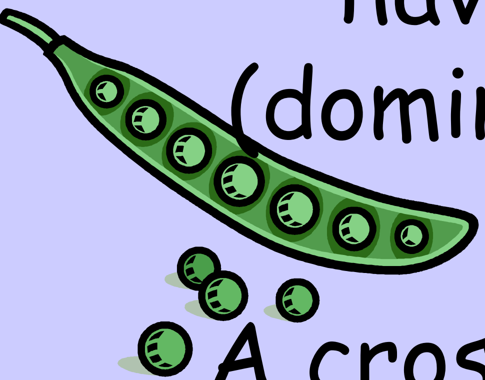
- the hypothesis to be tested is called the null hypothesis, Ho
- It is always stated that

  there is no difference between the observed and the expected results

- The formula is:-

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

example



- Pea plants may produce seeds that have round, yellow cotyledons (dominant) or wrinkled, green ones (recessive)
- A cross was made between a round, yellow plant and a wrinkled, green one and the offspring were then crossed giving the following results:

Phenotype of cotyledons	Number of offspring
round yellow	315
round green	108
wrinkled yellow	101
wrinkled green	32

If the parents are both heterozygous for both genes the expected offspring phenotype ratio would be

- 9 dominant for both characteristics
- 3 dominant for shape + recessive for colour
- 3 dominant for colour + recessive for shape
- 1 recessive for both characteristics

9:3:3:1

Is this the case???

Null hypothesis

there is no difference between the observed and the expected results

(i.e. the results fit the expected ratio)

Calculate the expected values if the ratio is true

Ratio =

- 9 dominant for both characteristics
- 3 dominant for shape + recessive for colour
- 3 dominant for colour + recessive for shape
- 1 recessive for both characteristics

The ratio adds up to 16

There are 556 offspring

- 312.75 dominant for both characteristics (9/16 of 556)
- 104.25 dominant for shape + recessive for colour (3/16 of 556)
- 104.25 dominant for colour + recessive for shape (3/16 of 556)
- 34.75 recessive for both characteristics (1/16 of 556)

Use the table format to calculate the value for χ^2

category	O	E	O-E	$(O-E)^2$	$\frac{(O-E)^2}{E}$
r/y	315	312.75	2.25	5.0625	0.0165187
r/g	108	104.25	3.75	14.0625	0.134892
w/y	101	104.25	3.25	10.5625	0.1013189
w/g	32	34.75	-2.75	7.5625	0.2176258
Σ	556	556			0.4703554 0.47

Degrees of freedom = number of categories -1

= n-1

Degrees of freedom = 4-1 **= 3**

An χ^2 value of 0.47

Use the χ^2 statistics table to find the 2 p values which the χ^2 value lies between...

Table 2: χ^2 values

d.f.	p = 0.900	0.500	0.100	0.050	0.010	0.001
1	0.016	0.455	2.71	3.84	6.63	10.83
2	0.211	1.39	4.61	5.99	9.21	13.82
3	0.584	2.37	6.25	7.81	11.34	16.27
4	1.06	3.36	7.78	9.49	13.28	18.47
5	1.61	4.35	9.24	11.07	15.09	20.52
6	2.20	5.35	10.64	12.59	16.81	22.46
7	2.83	6.35	12.02	14.07	18.48	24.32
8	3.49	7.34	13.36	15.51	20.09	26.13
9	4.17	8.34	14.68	16.92	21.67	27.88
10	4.87	9.34	15.99	18.31	23.21	29.59
11	5.58	10.34	17.28	19.68	24.73	31.26
12	6.30	11.34	18.55	21.03	26.22	32.91
13	7.04	12.34	19.81	22.36	27.69	34.53
14	7.79	13.34	21.06	23.68	29.14	36.12
15	8.55	14.34	22.31	25.00	30.58	37.70
16	9.31	15.34	23.54	26.30	32.00	39.25
17	10.09	16.34	24.77	27.59	33.41	40.79
18	10.86	17.34	25.99	28.87	34.81	42.31
19	11.65	18.34	27.20	30.14	36.19	43.82

If p is $<$ $p=0.05$ (the value you calculated for chi square is to the right on the table)

we must reject the null hypothesis

This means

there is a significant difference between the observed and expected results.

Any difference is NOT due to chance alone

If p is > p=0.05 (the value you calculated for chi square is to the right on the table)

we must accept the null hypothesis

This means any difference between

the observed and expected results

is due to chance alone

And is NOT significant.

i.e the observed results fit the expected 9:3:3:1 ratio

On the chi square table look at the row for 3 degrees of freedom.

The value 0.47 is less than the value of 0.584 at $p=0.9$.

$p=0.9$ is $>$ $p=0.05$.

Therefore accept null hypothesis

Difference due to chance alone and there is no significant difference between the observed and expected results

i.e the observed results fit the expected 9:3:3:1 ratio