Test for Significance : Chi-square test

⇒ Chi-square (χ^2) is a statical test which tell us the probability of obtaining the observed result by chance if the null hypothesis is true. It doesnot tell us whether the null hypothesis is ~~true~~ actually true.

⇒ The null hypothesis is that the data fits the ratio we suggest, and that what is tested

⇒ We must have a null hypothesis before we can do any meaningful test. our null hypothesis is that the results (63:37) do fit a 3:1.

⇒ $$\text{Chi-square }(\chi^2) = \text{Sum}\left[\frac{(\text{observed}-\text{expected})^2}{\text{Expected}}\right]$$

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

- χ^2 is calcullated from the original data, never from percentages, frequencies or proportions.

- χ^2 (chi-square) Calculations

Null hypothesis (Predicted ratio)	observed		
3:1	~~Expected~~ observed	63 Red	37 white
	Expected	$100 \times \frac{3}{4} = 75$	$100 \times \frac{1}{4} = 25$

- observed − expected = $63 - 75 = -12$ | $37 - 25 = 12$
- $(\text{observed}-\text{expected})^2 = (-12)^2 = 144.$ | $(12)^2 = 144.$
- $\frac{(\text{observed}-\text{expected})^2}{\text{expected}} = \frac{144}{75} = 1.92$ | $\frac{144}{25} = 5.76$

- $\chi^2 = 1.92 + 5.76 = 7.68.$
- degree of freedom $= 2-1 = 1$ d.f.
- from table of Probability (P)(d.f-1) $0.01 > P > 0.005.$

The expected numbers are calculateng by multiplying the total by the predicted frequency (eg. $180 \times \frac{1}{4}$ white). The deviation or difference, between the observed and expected number is squared to remove negative numbers and divided by the expected value to undo the squaring and standarize the numbers. This is repeated for each class (red & white classes in this example) and the values are added (Summed) to give the overall value of χ^2.
The value of χ^2 increases when the deviations from expected are large, so large values of χ^2 lead us to reject the null hypothesis.
- The data doesnot fit if χ^2 is large.
- A perfect fit gives a χ^2 of Zero.

Degree of freedom are one less than the number of classes.
- Degree of freedom are One less than the number of independent They ~~tell us~~ provide something about the number of independent numbers that have, and this relates to the usefulness of our data
- In the this example, there are two classes, red and white. All the plants that are not red must be white. when the Counted the red ones, the number of white one is fixed. Therefore, there is only one degree of freedom.
- If there had red, Pink and White flowers – three classes – there would be two degree of freedom, when two classes had been counted, the third would be fixed.

using χ^2 probability table

- first see the level of significance of our result.
- See the probability table and follow the line for 1 degree of freedom (top line) to find the nearest values of χ^2 above & below our value
- See that the value of 6.634 is exceeded with a Probability of 0.01 (that is 1% or $\frac{1}{100}$) and 7.87944 is exceeded with a probability of 0.005 (i.e. $0.5\% = \frac{1}{200}$).