



Chi Square (χ^2) Test

Testing for differences when the data are placed into categories
(nominal data)



When is it used?

- Tests differences between sets of data.
- Used with nominal data.
 - When data have been placed into categories.
 - The expected frequency should not be under 5 in more than 20% of the cells.



Question

- A student investigated whether the amount of lichen growing on a gravestone was influenced by the direction the sides face.
- She counted the number of lichen colonies on each of 3 sides and applied a chi square test.



State the hypotheses

- Null hypothesis:
 - There **is no** significant difference between the number of lichen colonies on each side of the gravestone.
- Alternative hypothesis:
 - There **is** a significant difference between the number of lichen colonies on each side of the gravestone.



Tabulate the data

	Number of colonies on the side facing...		
	North	South	West
Observed	21	33	54
Expected			

Calculate the expected frequencies if H_0 is true



	Number of colonies on the side facing...		
	North	South	West
Observed	21	33	54
Expected	36	36	36



Calculate the differences between observed & expected

	Number of colonies on the side facing...		
	North	South	West
Observed	21	33	54
Expected	36	36	36
O-E	-15	-3	18



Square these differences

	Number of colonies on the side facing...		
	North	South	West
Observed	21	33	54
Expected	36	36	36
O-E	-15	-3	18
$(O-E)^2$	225	9	324



Divide each square by the corresponding Expected frequency

	Number of colonies on the side facing...		
	North	South	West
Observed	21	33	54
Expected	36	36	36
O-E	-15	-3	18
$(O-E)^2$	225	9	324
$(O-E)^2/E$	6.25	0.25	9.00



Add together all of the $(O-E)^2/E$

$(O-E)^2/E$	6.25	0.25	9.00	15.50 (total)
-------------	------	------	------	-------------------------

$$\chi^2 = \sum (O-E)^2/E$$

$$\chi^2 = 15.5$$

Calculate the number of Degrees of Freedom

- For tables with one row of data,
 - DoF = number of columns – 1
- For tables containing more than one row of data,
 - DoF = (rows-1)x(columns-1)
- In this case, DoF = 3-1 = 2





Now look in printed χ^2 tables

- Look for the critical value (χ^2_{crit}) when
 - DoF = 2
 - Significance level is 0.05
- If your value for χ^2 is greater than the critical value, reject the null hypothesis & accept the alternative hypothesis.

df	P = 0.05	P = 0.01	P = 0.001
1	3.84	6.64	10.83
2	5.99	9.21	13.82
3	7.82	11.35	16.27
4	9.49	13.28	18.47
5	11.07	15.09	20.52
6	12.59	16.81	22.46
7	14.07	18.48	24.32
8	15.51	20.09	26.13
9	16.92	21.67	27.88
10	18.31	23.21	29.59



So...

- $\chi^2_{\text{crit}} = 5.99$
- $\chi^2 = 15.5$
- Reject H_0
- Accept H_A
 - There **is** a significant difference between the number of lichen colonies on each side of the gravestone.
 - We can say this with 95% certainty.