CHI - SQUARE TEST: for goodness of fit

$$\chi^2 = \Sigma \left[\frac{(O_i - E_i)^2}{E_i}\right]$$

where O_i = Observed frequency

E_i = Experimental (or) Expected frequency $= \frac{\Sigma O_i}{n}$

Degrees of freedom, $\nu = n-1$

Properties:

i) The mean of χ^2 distribution is equal to the no. of degrees of freedom

ii) The variance of χ^2 distribution is twice the degrees of freedom

iii) If χ^2 is a chi-square variate with ν degrees of freedom, then $\frac{\chi^2}{2}$ is a gamma variate with parameter $\nu/2$.

iv) Standard χ^2 variate tends to standard normal variate as $n \to \infty$

Applications:

i) To test if the hypothetical value of the population variance is $\sigma^2 = \sigma_0^2$.

ii) To test the goodness of fit.

iii) To test the independence of attributes.

iv) To test the homogenity of independent estimates of the population variance.

Degrees of freedom: No. of values in a set which may be assigned arbitrarily.

1. The table below gives the no. of aircraft accidents that occured during the various days of the week. Test whether the accidents are uniformly distributed over the week.

Days:	Mon	Tue	Wed	Thur	Fri	Sat
No. of accidents:	14	18	12	11	15	14

Total no. of accidents = 14 + 18 + 12 + 11 + 15 + 14 = 84

No. of days = 6.

$\therefore$ Expected frequency of accident $E_i = \frac{84}{6} = 14$.

Null hypothesis: H_0: The accidents are uniformly distributed

Alternative hypothesis: H_1: The accidents are not uniformly distributed

Level of significance: $\alpha = 5\%$ is fixed.

O_i	E_i	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
14	14	0	0
18	14	16	1.14
12	14	4	0.285
11	14	9	0.642
15	14	1	0.071
14	14	0	0
			2.14285

Test statistic: $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 2.14285$

Degrees of freedom: $\gamma = n - 1 = 6 - 1 = 5$.

$\chi^2_\alpha = 11.04$ at 5% with $\gamma = 5$

Conclusion: $\chi^2 = 2.14285 < 11.04 = \chi^2_\alpha$

$\therefore H_0$ is accepted at 5% level of significance.

(ie) The accidents are uniformly distributed.

2. A dice was thrown 498 times. Denoting x to be the no. appearing on the top face of it. The observed frequency of x is given below:

x: 1	2	3	4	5	6
f: 69	78	85	82	86	98

What option you could form for the accuracy of the dice?

$E_i = \frac{69 + 78 + 85 + 82 + 86 + 98}{6} = \frac{498}{6} = 83.$

O_i	E_i	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
69	83	196	2.3614
78	83	25	0.3012
85	83	4	0.0481
82	83	1	0.0120
86	83	9	0.1084
98	83	225	2.7108

$$\Sigma \frac{(O_i - E_i)^2}{E_i} \qquad 5.5419$$

Null hypothesis: H_0: A die is unbiased.

Alternative hypothesis: H_1: A die is not unbiased (ie) biased

Level of significance: $\alpha = 5\%$ is fixed

Test statistic: $\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i} = 5.5419$

Critical value: Degrees of freedom, $\gamma = n - 1 = 6 - 1 = 5$

$$\chi_\alpha^2 = 11.07$$

Conclusion: $\chi^2 = 5.5419 < 11.07 = \chi^2_\alpha$

$\therefore H_0$ is accepted at 5% level of significance

(ie) A die is unbiased.