CHI-SQUARE TEST (X2) :Test of goodness of fit

- Written as **Chi-square test** Pronounced as **Ki-Square**

- **X^2 =** Greek letter used by **Karl Pearson** in the year 1900.—to describe magnitude of difference between **Theory and Observation**.

- This test is mainly used in very **large sample size > 30**
- To test the accordance between fact and theory

- $$X^2 = \sum \left\{ \frac{(o-e)^2}{e} \right\}$$

- Where, O= observed freq. of the sample
- e= expected /calculated freq. of sample.

- This **X^2-** value gives idea about whether the difference between observed and expected freq. is **differ significantly or not**

CHI-SQUARE TEST REQUIREMENTS

1. Quantitative data.
2. One or more categories.
3. Independent observations.
4. Theoretical freq. should not be less than 5 (at least 10).
5. Simple random sample.
6. Data in frequency form.
7. All observations must be used.

Contd..

X^2- value tested with table value with d.df. (n-1).

n= frequencies of classes or Groups and not total observations.

Contingency Table and Yates Correction Factor:

Simple table with rows (**r**) and column (**c**):

Assume all cell freq. larger than 5

	Vaccinated	Not-Vaccinated	
Attacked	a	b	R1
Not-attacked	c	d	R2
	C1	C2	T

- **Ho=Cell freq. are independent**.

- **Estimation of expected freq. of respective cell**.

	Vaccinated	Not-Vaccinated
Attacked	$e = \frac{R1*C1}{T}$	$f = R1 - e$
Not-attacked	$g = C1 - e$	$h = C2 - f$ or $R2 - g$

- Therefore $X^2 = (a-e)^2/e + (b-f)^2/f + (c-g)^2/g + (d-h)^2/h$

- **D.F.** for Contingency table**=(r-1) (c-1)= (2-1)* (2-1)=1**
- **Conclusion: If calculated X^2– greater than table value of X^2 at 1df-- S**
- **If calculated X^2– less than table value of X^2 at 1df– NS**

Yates Correction Factor:

If freq. of any cell is **5 or less than 5,** Then correction factor is needed it was suggested by **F.YATES.**

Eg

2	8	10
18	**7**	25
20	15	T

Add 0.5 to each cell, whos diagonal cell product is less

subtract 0.5 from each cell- which Diagonal CP is more

These are corrected freq.

Further Procedure is same for Con. Table

‘F’ test : (‘F’ Distribution)

- Used to test the **homogeniety OR Equality of VARIANCES.**
- **It always estimated in terms of ratio of two variances of population keeping larger at the numerator, follows a distribution known as ‘F’ Distribution.**

- **Suppose:**
- **X_1, X_2----------Xn. Sample One with n1 sample size & $\bar{X}$-mean**
- **Y_1,Y_2 - - - - - - - - Yn. Sample two with n2 Sample size & $\bar{Y}$- mean.**

- **Population variance are**

- $S_1^2 = \sum (X_i - \bar{X})^2 / n_{1-1}$, $S_2^2 = \sum (Y - \bar{Y})^2 / n_{2-1}$.
- **Null Hypothesis Ho= Two samples have same variances for given population**

Contd..

-
-
- Test statistic= **F** = S_1^2 / S_2^2, with (n_1-1) and (n_2-1) df.
- i.e F. distribution has two degrees of freedoms.
- DF for numerator – given in 1st row
- Df for denominator -- given in 1st colmn.

In 'F' table.

- **Statement of Conclusion:**