CHI-square Test :-

→ The chi-square test for Goodness of fit a tests claims about population proportions [categorical variables]

→ It is a non-parametric test that is performed on categorical data. [ordinal, nominal data]

EX:- There is a population of Male who likes different color of bike

	Theory
yellow Bike	1/3
orange Bike	1/3
Red Bike	1/3

population proportions

⇓

Theoritical categorical distribution.

we will pick sample data. → In order to test the above claim

	Sample
yellow Bike	22
orange Bike	17
Red Bike	59

⇓

observed categorical distribution.

→ According to chi-square, with the help of sample data we need to come to conclusion on population proportions given. and this test is called Goodness of fit.

→ In 2010 census of the city, the weight of the individuals in a small city were found to be the following.

weight	<50 kg	50-75	>75
	20%	30%	50%

In 2020, weight of n=500 individuals were sampled. Below are the results.

<50	50-75	>75
140	160	200

using $\alpha = 0.05$, would you conclude the population differences of weights has changed in last 10 years?

Ans)

In 2010 Expected

<50 kg	50-75	>75
20%	30%	50%

In 2020 when, n=500 observed

<50	50-75	>75
140	160	200

In 2010 Expected

<50 kg	50-75	>75
500×0.2 = 100	500×0.3 = 150	500×0.5 = 250

① Null hypothesis (H_0): The data meets the expectation

Alternate hypothesis H_1: The data does not meet the expectation.

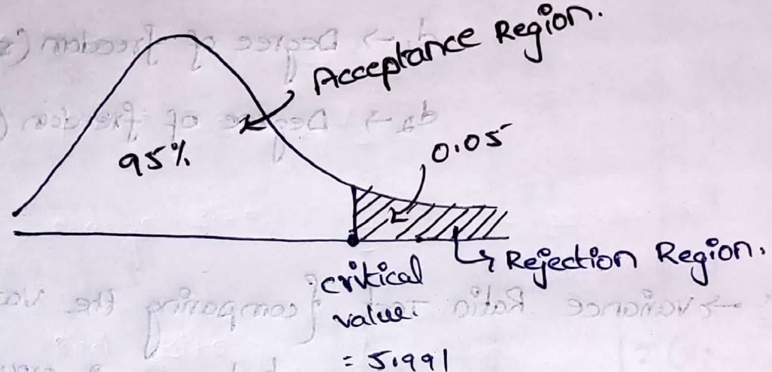
(2) $\alpha = 0.05$ $\quad C.I = 95\%$

(3) Degree of freedom

$dof = k-1 = 3-1 = 2$

(4) Decision Boundary:

→ usually chi-square distribution is a right-skewed Distribution.



chisquare Test $x^2 > 5.991$ {Reject the null hypothesis}

↳ From chi-square table.

(5) calculate chi square Test statistics

$$x^2 = \Sigma \frac{(O-E)^2}{E}$$

O – Observed

E – Expected.

$$= \frac{(140-100)^2}{100} + \frac{(160-150)^2}{150} + \frac{(200-250)^2}{250}$$

$$x^2 = 26.67$$

$x^2 > 5.991$ {Reject the Null hypothesis}

The weights of population has changed.

F-distribution :-

The F-distribution with d_1 and d_2 degrees of freedom is the distribution of

$$X = \frac{S_1/d_1}{S_2/d_2}$$

$S_1 \rightarrow$ Independent Random variables
$S_2 \rightarrow$ Independent Random variable
} chi square distribution

$d_1 \rightarrow$ Degree of freedom (S_1)

$d_2 \rightarrow$ Degree of freedom (S_2)

F-Test :- $\rightarrow$ Variance Ratio Test {comparing the Variance between 2 groups}

Q) The following data shows the no. of bulbs produced daily for some days by 2 workers A and B.

A	B
40	39
30	38
38	41
41	33
38	32
35	39
	40
	34

Can we consider based on the data worker B is more stable or not and efficient

$\alpha = 0.05$, 95% C.I

A) Null hypothesis $H_0 : \sigma_1^2 = \sigma_2^2$

Alternate hypothesis $H_1 : \sigma_1^2 \neq \sigma_2^2$

$n_1 = 6,\ n_2 = 8$

A

x_1	$\bar{x}$	$(x_i-\bar{x})^2$
40	37	9
30	37	49
38	37	1
41	37	16
38	37	1
35	37	4
$\bar{x}_1 = 37$		$\Sigma (x_i-\bar{x})^2 = 80$

B

x_2	$\bar{x}_2$	$(x_2-\bar{x})^2$
39	37	4
38	37	1
41	37	16
33	37	16
32	37	25
39	37	4
40	37	9
34	37	9
$\bar{x}_2 = 37$		$\Sigma (x_i-\bar{x}_2)^2 = 84$

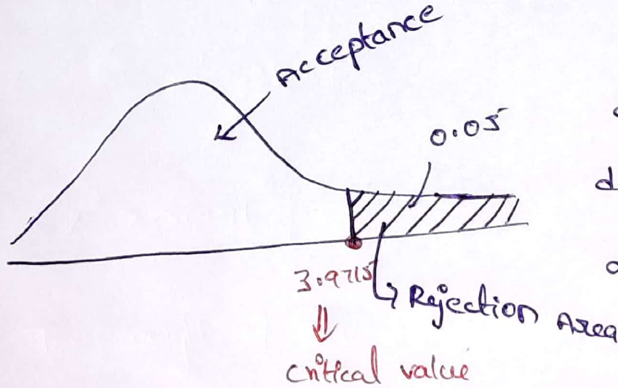
$$s_1^2 = \frac{\sum_{i=1}^{n} (x_i-\bar{x}_i)^2}{n-1}$$

$$= \frac{80}{5} = 16$$

* Variance Ratio [F-test]

$$F = \frac{s_1^2}{s_2^2} = \frac{16}{12} = 1.33$$

* Decision Rule :- [F-distribution] → is also a right-skewed distribution.



$df_1 = 6-1 = 5$

$df_2 = 8-1 = 7$

$\alpha = 0.05$

⇒ F-test > 3.9715 {Reject Null hypothesis}

Is 1.33 > 3.9715 ⇒ False.

we fail to reject the null hypothesis.

worker A ≈ worker B