Chi- square tests (X^2) :

The chi square tests are used to test the hypotheses and morals of nominal data, including:

1.Moral test for one sample (chi-square for quality of conciliation) .

2. Moral test for more than a sample (chi-square for independence) .

****

First: Moral test for one sample (chi square - quality of conciliation).

The chi square test uses the quality of conciliation to test whether the observed results differ from the expected results .

For the quality of conciliation: the conditions of conducting the test chi square .

1.The number of sample views is greater than 50 ($n > 50$).

2.The expected repetition corresponding to each category not less than 5 .

Chi square test steps for quality of conciliation : Formulation of Nullity and Alternative Hypothesis :

1. H_0 There is no difference between the observed results and the expected results .

2. H_1 There is a difference between the observed results and the expected results .

- The tablet value of the test statistic of chi square after the formation of a table helps us to calculate it as follows

Classes	f_o Actually Repetitions	f_e Expected Repetitions	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
Sum				$\sum\frac{(f_o - f_e)^2}{f_e}$	

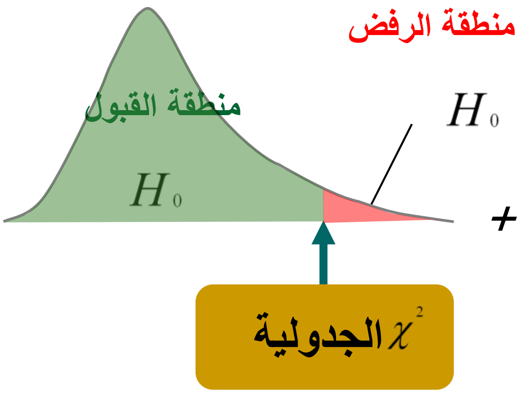
$$\chi^2 = \sum\frac{(f_o - f_e)^2}{f_e} :$$

The tablet value of chi square :

The determines the significant level by using degree of freedom from α (number of categories -1)

We extract the value of chi square : $\chi^2(n-1,\alpha)$.

4. Decision-making: We make the decision based on the value of the test statistic for the chi square (we specify the rejection area and acceptance area on the drawing following) :



H_1 We accept the alternative hypothesis if the value of the H_0 test falls in the rejection zone, we refuse to null hypothesis . and H_0 if the value of the test count in the area of acceptance, we accept the null hypothesis .

****

Example 1.In previous studies on mental patients, they were asked about their level of education

%5 at the undergraduate level :

%15 in secondary school :

%30 in the middle stage :

%50 at the primary level :

But currently the results for 60 persons are as follows :

Level of education	Number of patients
Undergraduate level	6
secondary school	20
middle stage	10
primary level	24
Sum	60

Can we decide that the actual results of this year's program are different from previous programs? (use $\alpha = 0.05$) .

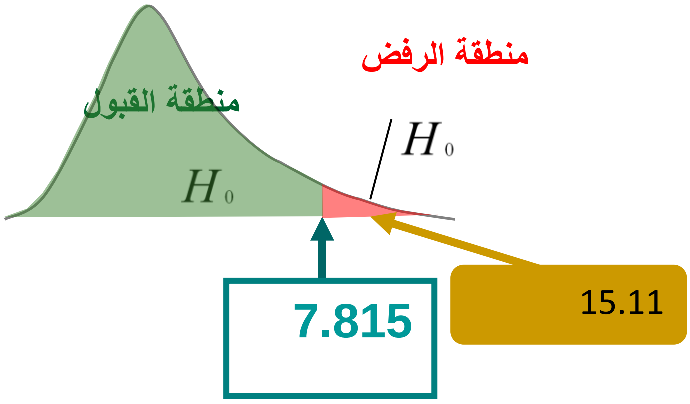
Solution:

1. H_0: There is no difference between the observed results and the expected results .
2. H_1: There is a difference between the observed results and the expected results .

change	f_o Observed repetitions	% percentage	f_e Expected repetitions	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
Undergraduate	**6**	**5%**	**0.05*60=3**	**3**	**9**	**3**
Secondary	**20**	**15%**	**0.15*60=11**	**9**	**81**	**7.36**
Intermediate	**10**	**30%**	**0.30*60=18**	**-8**	**64**	**3.55**
Primary	**24**	**50%**	**0.50*60=30**	**-6**	**36**	**1.2**
Sum	**60**					**15.11**

χ^2 calculated = 15.11 .,

χ^2 tablet = 7.815 .



Sign the test count in the rejection area :

We reject the null hypothesiss and accept alternative assumption that there is a difference between the observed results and the expected results .

Example 2. The Literacy Unit of the Ministry of Education has designed a propaganda program aimed at motivating and motivating the uneducated to change their attitudes so that they become more convinced of the benefit of education. The results of previous programs in this field are as follows:

%23 more believe in the importance of education (positive change) .

%65 do not change their direction (no change) .

12% change their attitudes to become more educated (negative change) .

For the year, the results of the program, which was conducted on 90 uneducated people, were as follows:

Change pattern	Number of persons
Positive change	52
No change	34
Negative change	4
Sum =	90

can we decide that the actual results of this year's program are different from previous programs? (use $\alpha = 0.05$) .

solution:

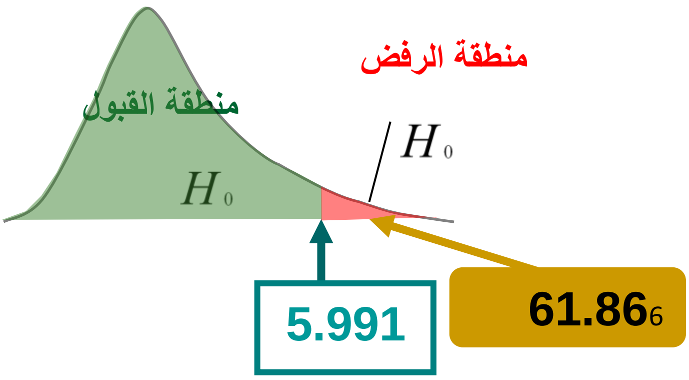
1. H_0: There is no difference between the observed results and the expected results .

2. H_1: There is a difference between the observed results and the expected results .

Change pattern	f_o Observed repetitions	% percentage	f_e Expected repetitions	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
Positive change	**52**	**23%**	**0.23*90=20.7**	**31.3**	**979.69**	**47.32**
No change	**34**	**65%**	**0.65*90=58.5**	**-24.5**	**600.25**	**10.26**
Negative change	**4**	**12%**	**0.12*90=10.8**	**-6.8**	**46.24**	**4.28**
Sum	**90**					**61.86**

χ^2 calculated = 61.86..,

χ^2 tablet = 5.991 .



the test count was lying in the rejection area .We refuse to null hypothesis and accept the alternative assumption that there is a difference between the observed results and the expected results

Second: Morality test for more than one sample (chi-square for independence).

In many cases, we need to identify whether there is a relationship between two classes of society. For example, we may need to know whether there is a relationship between income level and educational level. Or is there a relationship between eye color and hair color in a society? Or is there a relationship between the achievement level and household income? The chi- square for independence test is used to answer such questions (whether there is a relationship between two nominal variables or a nominal variable and an ordinal one) and is based on comparing observed values with expected values. Therefore, we should choose a random sample from the community

under study and then classify the observations of this sample according to the levels of each of the two characteristics and put them in a table called the compatibility table.

Steps of chi square test for independence : Formulation of Nullity and Alternative Hypothesis:

H_0: There is no relationship between the two qualities or there is no correlation between the two qualities

H_1: There is a relationship between the two characteristics or there is no correlation between the two qualities

2. the value of the test statistic chi-square : If both A and B have only two levels, the recurrence is a, b, c, and d, as follows:

	B_1	B_2
A_1	a	b
A_2	c	d

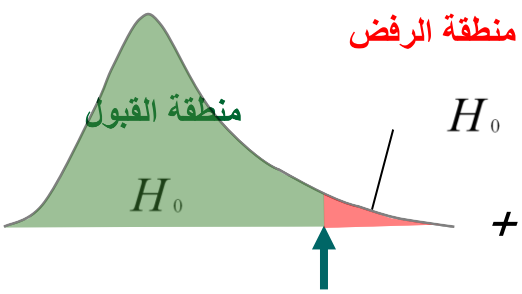
In this case, the test counts

$$\chi^2 = \frac{n(ad - bc)^2}{(a+b)(c+d)(a+c)(b+d)}$$

3. the tabular value of the chi square :

It has almost a square distribution of chi with one degree of freedom.

4. Decision-making: We make the decision based on the value of the test statistic .



H_1 We accept the alternative hypothesis if the H_0 value of the test statistic falls in the rejection zone, we refuse to null hypothesis .

H_0 But if the value of the test count in the area of acceptance, we accept the null hypothesis .

Example 3.: In a study to study the relationship between drinking tea and gender, a sample of 88 residents in a city was selected and classified in the following table. Does this data indicate a relationship between drinking tea and gender ? Use a significant level of $\alpha = 0.05$.

	Males	Females	Total
Drinking tea	40	33	73
No drinking tea	3	12	15
Sum	43	45	88

solution:

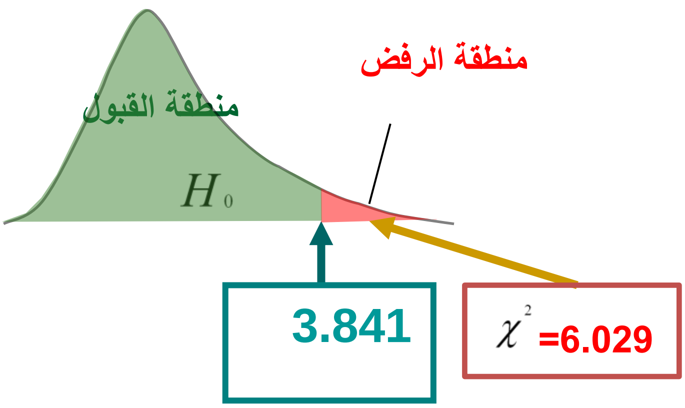
H_0 :There is no relationship between drinking tea and gender .

H_1 :There is a relationship between drinking tea and gender .

The value test statistic value is:

$$\chi^2 = \frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)} = \frac{88(480-99)^2}{73\times15\times43\times45} = 6.029$$

And we get the critical value of the distribution table of the square of chi square we find : $\chi^2(1,0.05) = 3.841$.



The value of the test statistic is greater than the tabular value, it is located in the rejection zone and therefore we reject H_0 and accept H_1 that there is a relationship between tea drinking and gender .

Example 4.: A social study was conducted to study the relationship between gender and the trend to marry relatives. A sample of 57 individuals was taken and the results were as follows ;

Gender trend to marry relatives.	Males	Females	Total
Pro-marriage	10	15	25
Unmarried	20	12	32
Total	30	27	57

Is there a correlation or relationship between gender and direction of relatives marriage or are the two independent of each other, which is not related between gender and direction of relatives marriage at a significant level 0.05 ? .

solution:

$: H_0$ There is no relationship between direction of relatives marriage and gender .

$: H_1$ There is a relationship between direction of relatives marriage and gender .

The test statistic value is : $\mathbf{X^2 = n\,(ad - bc)^2 / (a + b) \times (c + d) \times (a + c) \times (b + d)}$

$\mathbf{X^2 = 57\,(10 \times 12 - 15 \times 20)^2 / (10 + 15) \times (20 + 12) \times (10 + 20) \times (15 + 12) =}$

$\mathbf{X^2 = 57\,(120 - 300)^2 / (25) \times (32) \times (30) \times (27) =}$

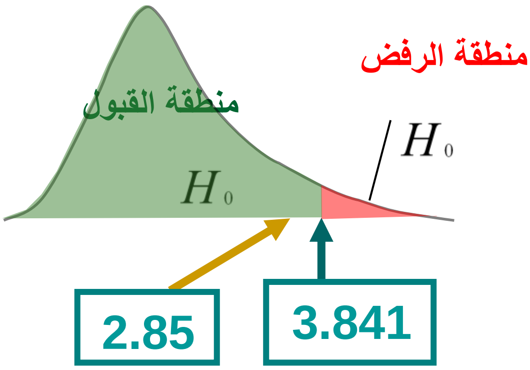
$\mathbf{X^2 = 57\,(-180)^2 / 648000 =}$

$\mathbf{X^2 = 57\,(32400) / 648000 =}$

$\mathbf{X^2 = 1846800 / 648000 =}$

$\mathbf{X^2 = 2.85}$ **.**

And we get the critical value of the distribution table of the chi square we find: $\chi^2(1,0.05) = 3.841$ The value of the test statistic is greater than the tabular value, it is located in the rejection zone and therefore we reject H_0 and accept H_1 that there is a relationship between tea drinking and gender .



The value of the test statistic is smaller than the tabular value, it falls into the acceptance area and therefore we accept null hypothesis and refuse the alternative hypothesis so that there is no relationship between the direction of relatives marriage and gender .