Chi-square test

Chi-square Distribution $\chi^2(\nu)$

Definition:

The pdf is,

$$f(\chi^2) = \frac{1}{2^{\nu/2}\sqrt{\left(\frac{\nu}{2}\right)}} \cdot (\chi^2)^{\nu/2-1} e^{-\chi^2/2}$$

$0 < \chi^2 < \infty$, where ν is the number of degrees of freedom.

Properties of χ^2-Distribution

1. A rough sketch of the probability curve of the χ^2-distribution for $\nu = 3$ and $\nu = 6$ is given in Fig.
2. As ν becomes smaller and smaller, the curve is skewed more and more to the right. As ν increases, the curve becomes more and more symmetrical.
3. The mean and variance of the χ^2-distribution are ν and 2ν respectively.

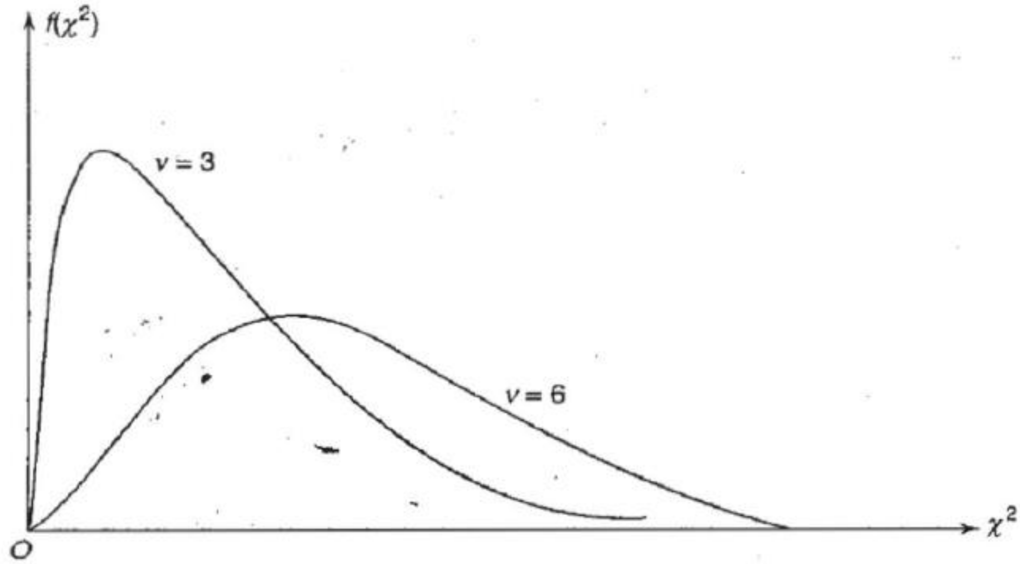


Fig.

4. As n tends to ∞, the χ^2-distribution becomes a normal distribution.

Uses of χ^2-Distribution

1. χ^2-distribution is used to test the goodness of fit. i.e., it is used to judge whether a given sample may be reasonably regarded as *a* simple sample from a certain hypothetical population.
2. It is used to test the independence of attributes. i.e. If a population is known to have two attributes (or traits), then χ^2-distribution is used to test whether the two attributes are associated or independent, based on a sample drawn from the population.

■ II. Application or uses of χ^2-distribution ■

(1) To test the "goodness of fit".

(2) To test the "independence of attributes".

(3) To test if the hypothetical value of the population variance is σ

(4) To test the homogeneity of independent estimates of the population variance.

(5) To test the homogeneity of independent estimates of the population correlation coefficient.

χ^2-Test of Goodness of Fit

On the basis of the hypothesis assumed about the population, we find the expected frequencies $E_i (i = 1, 2, ..., n)$, corresponding to the observed frequencies $O_i (i = 1, 2, ..., n)$ such that $\boxed{\sum E_i = \sum O_i}$. It is known that $\boxed{\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}}$ follows approximately a χ^2-distribution with degrees of freedom equal to the number of independent frequencies. In order to test the goodness of fit, we have to determine how far the differences between O_i and E_i can be attributed to fluctuations of sampling and when we can assert that the differences are large enough to conclude that the sample is not a simple sample from the hypothetical population. In other words, we have to determine how large a value of χ^2 we can get so as to assume that the sample is a simple sample from the hypothetical population.

Note: $\nu = n - 1$

If the calculated $\chi^2 < \chi^2_\nu(\alpha)$, we will accept the null hypothesis H_0 which assumes that the given sample is one drawn from the hypothetical population, i.e. we will conclude that the difference between the observed and expected frequencies is not significant at α % LOS If $\chi^2 > \chi^2_\nu(\alpha)$, we will reject H_0 and conclude that the difference is significant.

Conditions for the Validity of χ^2-Test

1. The number of observations N in the sample must be reasonably large, say ≥ 50.
2. Individual frequencies must not be too small, i.e. $O_i \geq 10$. In case $O_i < 10$, it is combined with the neighbouring frequencies, so that the combined frequency is ≥ 10.
3. The number of classes n must be neither too small nor too large i.e., $4 \leq n \leq 16$.

Book-2

1.3.b. χ^2-test to test the goodness of fit.

A very powerful test for testing the significance of the discrepancy between theory and experiment was given by prof. Karl-Pearson in 1990 and is known as "Chi-square test of goodness of fit". It enables us to find if the deviation of the experiment from theory is just by chance or is it really due to the inadequacy of the theory to fit the observed data.

■ I. Chi-Square Test for Goodness of fit ■

χ^2-test of goodness of fit is a test to find if the deviation of the experiment from theory is just by chance or it is due to the inadequacy of the theory to fit the observed data.

By this test, we test whether differences between observed and expected frequencies are significant or not.

χ^2-test statistic of goodness of fit is defined by

$$\chi^2 = \Sigma \frac{(O - E)^2}{E}, \text{ where } \quad O \rightarrow \text{Observed frequency}$$

$$E \rightarrow \text{Expected frequency}$$

■ III. Conditions for the application of χ^2-test. ■

[AU N/D 2011]

(1) The sample observations should be independent.

(2) Constraints on the cell frequencies, if any, must be linear [e.g., $\Sigma O_i = \Sigma E_i$]

(3) N, the total frequency, should be atleast 50.

(4) No theoretical cell frequency should be less than 5.

■ IV. Independence of attributes ■

Note 1 In the case of

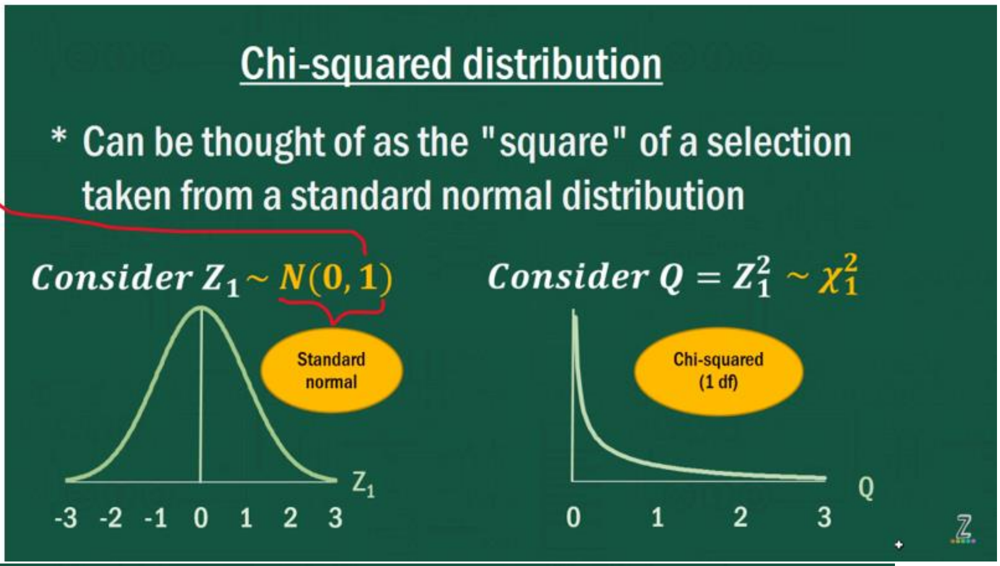
fitting a Binomial distribution, d.f $= n - 1$

fitting a Poisson distribution, d.f $= n - 2$

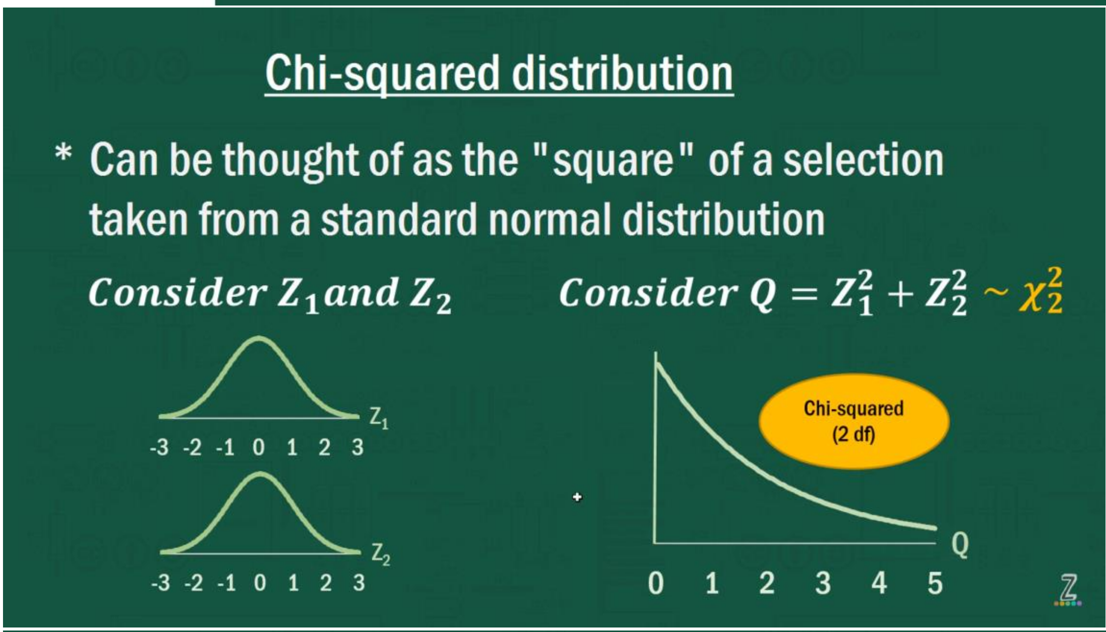
fitting a Normal distribution, d.f $= n - 3$

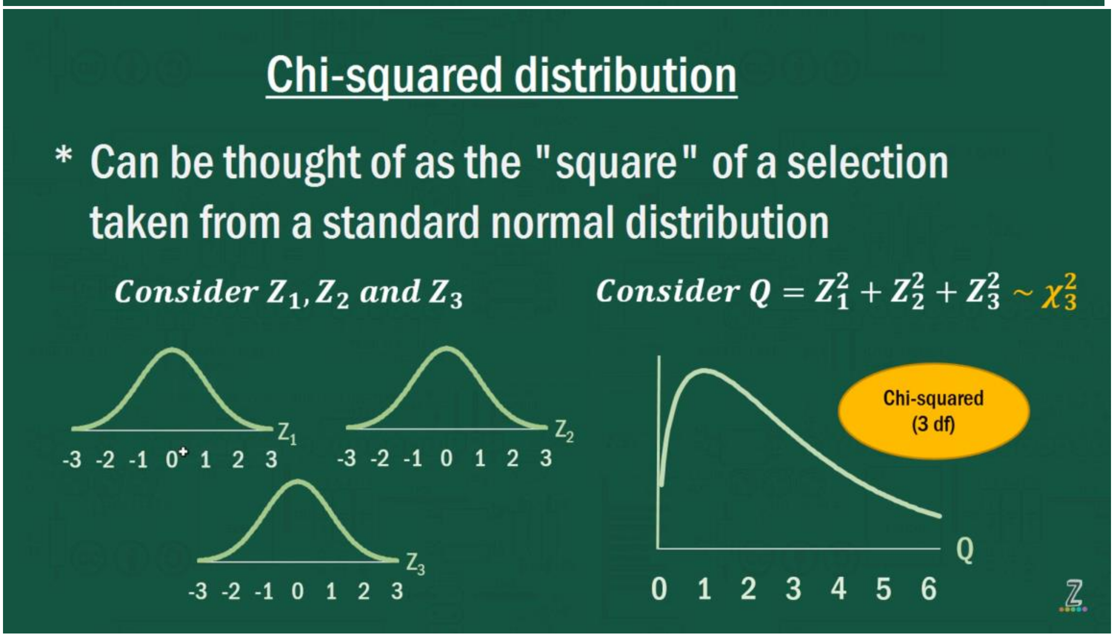
Note 2 If $\chi^2 = 0$, all observed and expected frequencies coincide

Note 3 For χ^2-distribution, mean $= \nu$, Variance $= 2\nu$

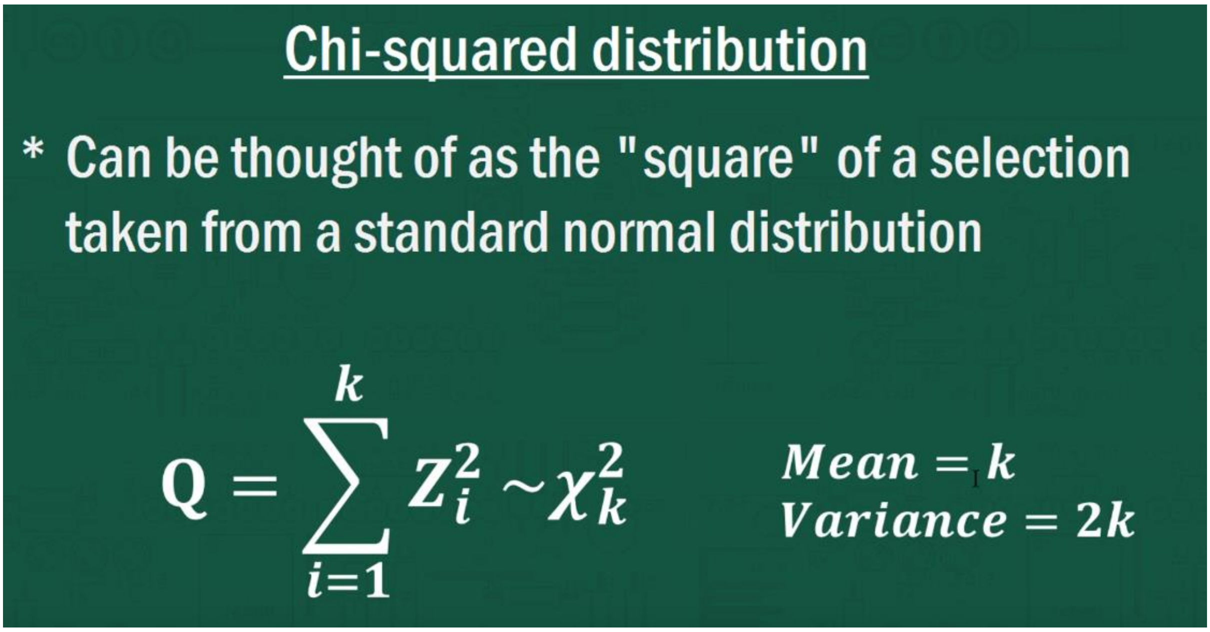


variance of dist

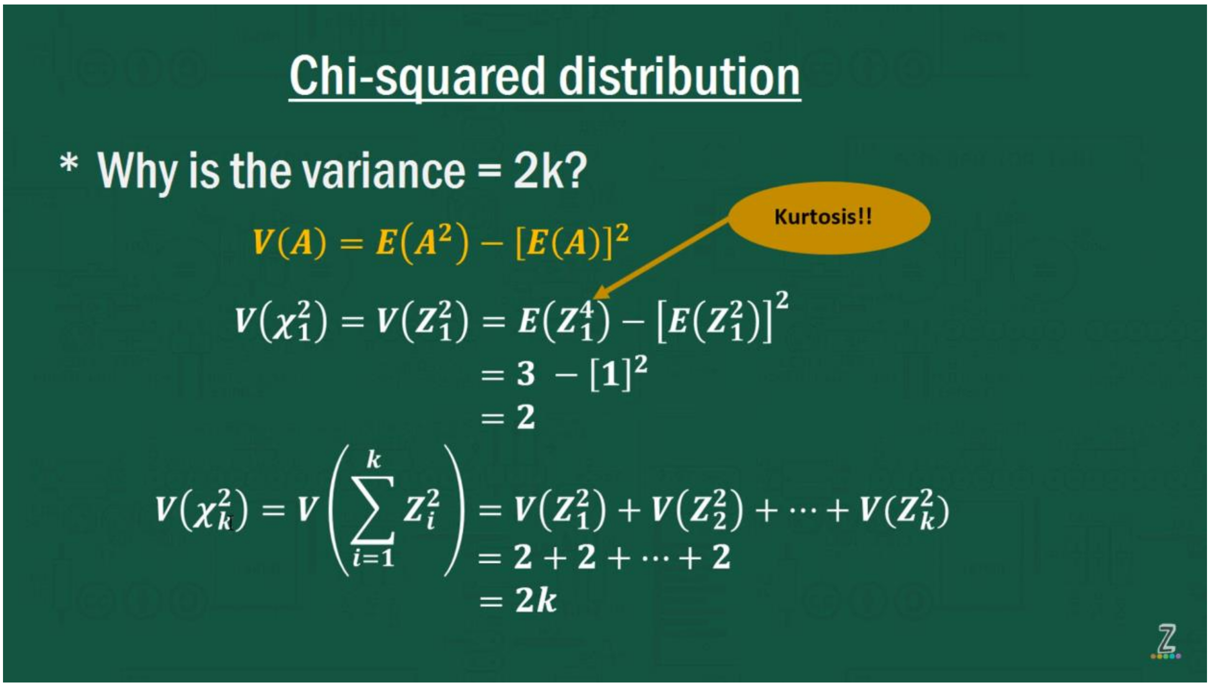




Why mean=K (with K degrees of freedom) ????



Why variance=2K (with K degrees of freedom) ????



Question: (N=256 and n=6) (0^{th} is included)

Five coins are tossed 256 times. The number of heads observed is given below. Examine if the coins are unbiased, by employing χ^2 goodness of fit.

No. of heads	0	1	2	3	4	5
Frequency	5	35	75	84	45	12

Solution :

Given : $n = 6$, N = total number of frequencies = 256

1. H_0 : Binomial is a good fit.
2. H_1 : Binomial is not a good fit.
3. $\alpha = 0.05$, d.f $= n - 1 = 6 - 1 = 5$
4. Table value of $\chi^2 = 11.07$
5. **The test statistic is** $\chi^2 = \Sigma \frac{(O-E)^2}{E}$

On the assumption H_0, the expected frequencies are given by the terms of $N\,(q+p)^n = 256\left(\frac{1}{2}+\frac{1}{2}\right)^5$

$$= \frac{256}{32}[5c_0 + 5c_1 + 5c_2 + 5c_3 + 5c_4 + 5c_5]$$

$$= \frac{256}{32}[1 + 5 + 10 + 10 + 5 + 1]$$

$$= 8\,[1 + 5 + 10 + 10 + 5 + 1]$$

$\therefore$ The expected frequencies are

8, 40, 80, 80, 40, 8

No. of heads	O	E	O − E	$(O-E)^2$	$\frac{(O-E)^2}{E}$
0	5	8	−3	9	1.2500
1	35	40	−5	25	0.6250
2	75	80	−5	25	0.3125
3	84	80	4	16	0.2000
4	45	40	5	25	0.6250
5	12	8	4	16	2.000
	256	256			4.8875

$$\chi^2 = \Sigma \frac{(O-E)^2}{E} = 4.8875$$

6. Conclusion :

If cal. $\chi^2 <$ table χ^2, then we accept H_0 ; otherwise, we reject H_0.

Here, $\chi^2 = 4.8875 < 11.07$. So, we accept to at 5% level of signiciance.

$\therefore$ Binomial distribution is a good fit to the given data.

Question: **(N=160 and n=5) (0^{th} is included)**

4 coins were tossed 160 times and the following results were obtained:

No. of heads :	0	1	2	3	4
Observed frequencies :	17	52	54	31	6

Under the assumption that the coins are unbiased, find the expected frequencies of getting 0, 1, 2, 3, 4 heads and test the goodness of fit.

Solution : **[AU A/M 2011]**

1. **Null hypothesis H_0** : The coins are unbiased
2. **Alternative hypothesis H_1** : The coins are biased
3. **Level of significance** : $\alpha = 0.05$, d.f. $= n - 1 = 5 - 1 = 4$
4. **Table value of χ^2** : 9.488
5. **Test statistic** : Under H_0, the test statistic is

$$\chi^2 = \sum \frac{(0-E)^2}{E} \sim \chi^2 \text{ distribution with } n-1 \text{ d.f.}$$

$$\text{Probability of getting head} = p = \frac{1}{2}$$

$$\text{Probability of getting tail} = q = \frac{1}{2}$$

Then the expected frequencies are

$$p(x) = n\,C_x p^x q^{n-x},\ x = 0, 1, 2, 3, \dots$$

$$p(0 \text{ head}) = 4\,C_0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^4 = 0.0625$$

$$p(1 \text{ head}) = 0.25\ ,\quad p\ (2 \text{ heads}) = 0.375$$

$$p(3 \text{ head}) = 0.25\ ,\quad p\ (4 \text{ heads}) = 0.0625$$

χ^2 value is calculated from the following table :

No. of heads (x_i)	O	p (x_i)	E = (160 × p (x_i))	$\frac{(O-E)^2}{E}$
0	17	0.0625	10	4.9
1	52	0.25	40	3.6
2	54	0.375	60	0.6
3	31	0.25	40	2.025
4	6	0.0625	10	1.6
Total	160		160	12.725

Calculated χ^2 = 12.725

6. Conclusion :

If cal. χ^2 < table χ^2, then we accept H_0 ; otherwise, we reject H_0.

Here, 12.725 $\nless$ 9.488

So, we reject H_0 and accept H_1

i.e., the coins are biased.

$N(p+q)^n$

$\frac{160}{16}\left[{}^4C_0 + {}^4C_1 + {}^4C_2 + {}^4C_3 + {}^4C_4\right]$

$\boxed{160\left(\frac{1}{2}+\frac{1}{2}\right)^4}$

$10\,[1 + 4 + 6 + 4 + 1]$

$\Rightarrow\ 10 + 40 + 60 + 40 + 10$

Question **(very very important)** **(N=90 and n=6)**

Example:

Table gives the number of air-craft accidents that occurred during the various days of a week. Test whether the accidents are uniformly distributed over the week.

Table

Day:	Mon	Tues	Wed	Thu	Fri	Sat	Total
No. of accidents:	15	19	13	12	16	15	90

Soln/ $N = 90$ & $n = 6$.

①. H_0: The accidents are distributed uniformly.

②. H_1: Not uniformly distributed.

③. $LOS = \alpha = 5\% = 0.05$

$$\chi^2_{Tab} = \chi^2_{5\%}(\nu = 5) = ?$$

④. Test Statistic:

$$\chi^2_{Cal} = \sum_{i=1}^{6} \frac{(O_i - E_i)^2}{E_i}$$

$E_i = \frac{90}{6} = 15$; $i = 1, 2, \ldots, 6$

[Since uniform distribution]

⑤. Comparison and Conclusion

Solution:

H_0 : Accidents occur uniformly over the week.

Total number of accidents = 90

Based on H_0, the expected number of accidents on any day $= \dfrac{90}{6} = 15.$

O_i	:	15	19	13	12	16	15
E_i	:	15	15	15	15	15	15

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = \frac{1}{15}(0 + 16 + 4 + 9 + 1 + 0) = 2.$$

Since $\sum E_i = \sum O_i, \quad v = 6 - 1 = 5$

From the χ^2-table, $\chi^2_{5\%}(v = 5) = 11.07.$

Since $\chi^2 < \chi^2_{5\%}$, H_0 is accepted.

i.e. accidents may be regarded to occur uniformly over the week.

Question:

Example:

Theory predicts that the proportion of beans in four groups A, B, C, D should be 9 : 3 : 3 : 1. In an experiment among 1600 beans, the numbers in the four groups were 882, 313, 287 and 118. Does the experiment support the theory?

Solution:

H_0 : The experiment supports the theory, i.e. the numbers of beans in the four groups are in the ratio 9 : 3 : 3 : 1

Based on H_0, the expected numbers of beans in the four groups are as follows

$$E_i : \quad \frac{9}{16} \times 1600, \quad \frac{3}{16} \times 1600, \quad \frac{3}{16} \times 1600, \quad \frac{1}{16} \times 1600$$

i.e.	E_i :	900,	300,	300,	100
	O_i :	882,	313,	287,	118

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = \frac{18^2}{900} + \frac{13^2}{300} + \frac{13^2}{300} + \frac{18^2}{100} = 4.73$$

Since $\sum E_i = \sum O_i$, $v = 4 - 1 = 3$

From the χ^2-table, $\chi^2_{5\%}(v = 3) = 7.82$

Since $\chi^2 < \chi^2_{5\%}$, H_0 is accepted.

i.e. the experimental data support the theory.

Question: **very very important** **(N=80 and n=6) (0^{th} is not included)**

Example:

Fit a binomial distribution for the following data and also test the goodness of fit.

x:	0	1	2	3	4	5	6	Total
f:	5	18	28	12	7	6	4	80

Solution:

To find the binomial distribution $N(q+p)^n$, which fits the given data, we require p.

We know that the mean of the binomial distribution is np, from which we can find p. Now the mean of the given distribution is found out and is equated to np.

x:	0	1	2	3	4	5	6	Total
f:	5	18	28	12	7	6	4	80
fx:	0	18	56	36	28	30	24	192

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{192}{80} = 2.4$$

i.e. $np = 2.4$ or $6p = 2.4$, since the maximum value taken by x is n.

$\therefore$ $p = 0.4$ and hence $q = 0.6$

$\therefore$ The expected frequencies are given by the successive terms in the expansion of $80\,(0.6+0.4)^6$.

Thus E_i: 3.73, 14.93, 24.88, 22.12, 11.06, 2.95, 0.33

Converting the E_i's into whole number such that $\sum E_i = \sum O_i = 80$, we get

E_i: 4 15 25 22 11 3 0

Let us now proceed to test the goodness of binomial fit.

O_i: 5 18 28 12 7 6 4

$E_i > 5$

The first class is combined with the second and the last two classes are combined with the last but second class in order to make the expected frequency in each class greater than or equal to 10. Thus, after regrouping, we have,

E_i:	19	25	22	14
O_i:	23	28	12	17

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = \frac{4^2}{19} + \frac{3^2}{25} + \frac{10^2}{22} + \frac{3^2}{14} = 6.39$$

We have used the given sample to find

$\sum E_i \; (= \sum O_i)$ and p through its mean.

Hence $\nu = n - k$

$= 4 - 2 = 2$

$\chi^2_{5\%}\,(\nu = 2) = 5.99$, from the χ^2-table.

Since $\chi^2 > \chi^2_{5\%}$, H_0, which assumes that the given distribution is approximately a binomial distribution, is rejected. i.e. the binomial fit for the given distribution is not satisfactory.

Question: **very very important** **(N=47 and n=6) (0^{th} is included)**

Example 5. *Verify whether the Poisson distribution can be assumed from the data given below:*

No. of defects	*0*	*1*	*2*	*3*	*4*	*5*
Frequency	*6*	*13*	*13*	*8*	*4*	*3*

Sol. H_0 : The Poisson fit is a good fit to the data.

Mean of the given distribution $= \dfrac{\Sigma f_i x_i}{\Sigma f_i} = \dfrac{94}{47} = 2$

To fit a Poisson distribution we require m. Parameter $m = \bar{x} = 2$.
By the Poisson distribution the frequency of r success is

$$N(r) = N \times e^{-m} \cdot \frac{m^r}{r!}, \text{ N is the total frequency.}$$

$$N(0) = 47 \times e^{-2} \cdot \frac{(2)^0}{0!} = 6.36 \approx 6; \qquad N(1) = 47 \times e^{-2} \cdot \frac{(2)^1}{1!} = 12.72 \approx 13$$

$$N(2) = 47 \times e^{-2} \cdot \frac{(2)^2}{2!} = 12.72 \approx 13; \qquad N(3) = 47 \times e^{-2} \cdot \frac{(2)^3}{3!} = 8.48 \approx 9$$

$$N(4) = 47 \times e^{-2} \cdot \frac{(2)^4}{4!} = 4.24 \approx 4; \qquad N(5) = 47 \times e^{-2} \cdot \frac{(2)^5}{5!} = 1.696 \approx 2.$$

X	0	1	2	3	4	5
O_i	6	13	13	8	4	3
E_i	6.36	12.72	12.72	8.48	4.24	1.696
$\dfrac{(O_i - E_i)^2}{E_i}$	0.2037	0.00616	0.00616	0.02716	0.0135	1.0026

$$\chi^2 = \frac{\Sigma (O_i - E_i)^2}{E_i} = 1.2864.$$

Conclusion. The calculated value of χ^2 is 1.2864. The tabulated value of χ^2 at 5% level of significance for $\gamma = 6 - 2 = 4$ d.f. is 9.49. Since the calculated value of χ^2 is less than that of the tabulated value, H_0 is accepted, *i.e.*, the Poisson distribution provides a good fit to the data.

Question(N=90 and n=10) (0^{th} is included)

EXAMPLE 24.107

The following table gives the frequency of occupancy of digits 0, 1, 2,…,9 in the last place in four logarithms of numbers 10–99. Examine if there is any peculiarity.

n => 10

Digits:	0	1	2	3	4	5	6	7	8	9	
Frequency:	6	16	15	10	12	12	3	2	9	5	90

Solution. Let the null hypothesis be
H_0: frequency of occupance of digits is equal, that is, there is no significant difference between the observed and the expected frequency.

Therefore under the null hypothesis, the expected frequency is $f_e = \frac{90}{10} = 9.$ Then

$$\chi^2 = \frac{\sum (f_{o_i} - f_{e_i})^2}{f_{e_i}}$$

$$= \frac{9 + 49 + 36 + 1 + 9 + 9 + 36 + 49 + 0 + 16}{9}$$

$$= 23.777.$$

Number of degree of freedom is 10 –1 = 9. The tabulated value of $\chi^2_{0.05}$ for $v = 9$ is 16.92. Since the calculated value of χ^2 is greater than the tabulated value of $\chi^2_{0.05}$, the hypothesis is rejected and so there is a significant difference between the observed and expected frequency.

Question(N=10000 and n=10) (0^{th} is included)

The following table shows the distribution of digits in the numbers chosen at random from a telephone directory:

Table

Digit:	0	1	2	3	4	5	6	7	8	9	Total
Frequency:	1026,	1107,	997,	966,	1075	933,	1107,	972,	964,	853	10,000

Test whether the digits may be taken to occur equally frequently in the directory.

Sol. Null hypothesis H_0: The digits taken in the directory occur equally frequently.

i.e., there is no significant difference between the observed and expected frequency.

Under H_0, the expected frequency is given by $= \dfrac{10,000}{10} = 1000$

To find the value of χ^2

O_i	1026	1107	997	996	1075	1107	933	972	964	853
E_i	1000	1000	1000	1000	1000	1000	1107	1000	1000	1000
$(O_i - E_i)^2$	676	11449	9	1156	5625	11449	4489	784	1296	21609

$$\chi^2 = \frac{\Sigma(O_i - E_i)^2}{E_i} = \frac{58542}{1000} = 58.542.$$

Conclusion. The tabulated value of χ^2 at 5% level of significance for 9 difference is 16.919. Since the calculated value of χ^2 is greater than the tabulated value, H_0 is rejected.

There is significant difference between the observed and theoretical frequency.

The digits taken in the directory do not occur equally frequently.

Question very very important (N=800 and n=5) (0th is included)

Example 4. *Records taken of the number of male and female births in 800 families having four children are as follows:*

No. of male births	*0*	*1*	*2*	*3*	*4*
No. of female births	*4*	*3*	*2*	*1*	*0*
No. of families	*32*	*178*	*290*	*236*	*94*

Test whether the data are consistent with the hypothesis that the binomial law holds and the chance of male birth is equal to that of female birth, namely $p = q = 1/2$.

Sol. H_0: The data are consistent with the hypothesis of equal probability for male and female births, *i.e.*, $p = q = 1/2$.

We use binomial distribution to calculate theoretical frequency given by:

$$N(r) = N \times P(X = r)$$

where N is the total frequency. $N(r)$ is the number of families with r male children:

$$P(X = r) = {}^{n}C_r p^r q^{n-r}$$

where p and q are the probability of male and female births, n is the number of children.

$$N(0) = \text{No. of families with 0 male children} = 800 \times {}^4C_0 \left(\frac{1}{2}\right)^4 = 800 \times 1 \times \frac{1}{2^4} = 50$$

$$N(1) = 800 \times {}^4C_1 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^3 = 200; \quad N(2) = 800 \times {}^4C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^2 = 300$$

$$N(3) = 800 \times {}^4C_3 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^3 = 200; \quad N(4) = 800 \times {}^4C_4 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^4 = 50$$

Observed frequency O_i	32	178	290	236	94
Expected frequency E_i	50	200	300	200	50
$(O_i - E_i)^2$	324	484	100	1296	1936
$\frac{(O_i - E_i)^2}{E_i}$	6.48	2.42	0.333	6.48	38.72

$$\chi^2 = \frac{\Sigma(O_i - E_i)^2}{E_i} = 54.433.$$

Conclusion. The table value of χ^2 at 5% level of significance for $5 - 1 = 4$ d.f. is 9.49.
Since the calculated value of χ^2 is greater than the tabulated value, H_0 is rejected.
I.e., the data are not consistent with the hypothesis that the binomial law holds and that the chance of a male birth is not equal to that of a female birth.

Note. Since the fitting is binomial, the degrees of freedom $\nu = n - 1$, *i.e.*, $\nu = 5 - 1 = 4$.

Question

3. A survey of 320 families with 5 children each revealed the following information:

No. of boys	:	5	4	3	2	1	0
No. of girls	:	0	1	2	3	4	5
No. of families	:	14	56	110	88	40	12

Is this result consistent with the hypothesis that male and female births are equally probable?

Question

Fit a Poisson distribution to the following data and the best goodness of fit:

x	:	0	1	2	3	4
f	:	109	65	22	3	1

Question

The sales in a supermarket during a week are given below. Test the hypothesis that the sales do not depend on the day of the week, using a significance level of 0.05.

Days	:	Mon	Tues	Wed	Thurs	Fri	Sat
Sales (in $10000)	:	65	54	60	56	71	84

Question

In the accounting department of a bank, 100 accounts are selected at random and estimated for errors. The following results were obtained:

No. of errors	:	0	1	2	3	4	5	6
No. of accounts	:	35	40	19	2	0	2	2

Does this information verify that the errors are distributed according to the Poisson probability law?

Number of errors	0	1	2	3	4	5	6
Number of accounts	35	40	19	2	0	2	2

$\boxed{N \Rightarrow 100}$

$n \Rightarrow 100$ accounts

$$\text{Mean} \Rightarrow \lambda \Rightarrow \frac{\Sigma fx}{\Sigma f} \Rightarrow \frac{40+38+6+10+12}{35+40+19+6} \Rightarrow \frac{106}{100} \Rightarrow 1.06$$

$\boxed{\lambda \Rightarrow 1.06}$

$\boxed{n \Rightarrow 7}$

$df \Rightarrow n-2$

$\Rightarrow 7-2 \Rightarrow 5$

$\boxed{df \Rightarrow 5}$

Poisson distribution $p(x) \Rightarrow \dfrac{e^{-\lambda}\lambda^x}{x!}$

$$p(0) \Rightarrow \frac{e^{-1.06}(1.06)^0}{0!} \Rightarrow 0.346 \Rightarrow \boxed{34.6}$$

$$p(1) \Rightarrow \frac{e^{-1.06}(1.06)^1}{1!} \Rightarrow 0.3672 \Rightarrow \boxed{36.7}$$

$$p(2) \Rightarrow \frac{e^{-1.06}(1.06)^2}{2!} \Rightarrow 0.1946 \Rightarrow \boxed{19.4}$$

$$p(3) \Rightarrow \frac{e^{-1.06}(1.06)^3}{3!} \Rightarrow 0.06877 \quad \boxed{6.8}$$

$$p(4) \Rightarrow \frac{e^{-1.06}(1.06)^4}{4!} \Rightarrow 0.1518 \quad \boxed{15.1}$$

$$p(5) \Rightarrow \frac{e^{-1.06}(1.06)^5}{5!}$$

$$p(6) \Rightarrow 0.003863 \quad \boxed{0.38}$$

$$p(6) \Rightarrow \frac{e^{-1.06}(1.06)^6}{6!}$$

$\Rightarrow 0.0006825$ $\boxed{0.068}$

O	E	$(O-E)^2$	$(O-E)^2/E$
35	34.6	0.16	0.004624
40	36.7	10.89	0.2967
19	19.4	0.16	0.008247
2	6.8	23.04	3.388
0	15.1	228.01	15.1
2	0.38	2.6244	2.2444
2	0.068	3.732	54.88

$$\chi^2 \Rightarrow \Sigma\left((O-E)^2/E\right)$$

$\boxed{\chi^2 \Rightarrow 75.921}$

$\boxed{\chi^2_\alpha \Rightarrow 11.070}$

$\chi^2_\alpha > \chi^2$ H_0 is accepted

Question

A sample analysis of examination results of 500 students was made. It was found that 220 students had failed, 170 had secured a third class, 90 were placed in second class and 20 got a first class. Do these figures commensurate with the general examination result which is in the ratio of 4: 3: 2: 1 for the various categories respectively.

Solution:

Null hypothesis H_0 : The observed results commensurate with the general examination results.

$N \Rightarrow 500$ student

500 → 280, 220 fail (4)

280:

(3) ↳ 170 ⇒ C's

(2) ↳ 90 ⇒ B's

(1) ↳ 20 ⇒ A's

Ratio ⇒ 4 : 3 : 2 : 1

(4+3+2+1 ⇒ 10)

Class	First	Second	Third	Fail
Observed frequency (O)	20	90	170	220
Expected frequency (E)	50	100	150	200
$(O-E)^2$	900	100	400	400
$\frac{(O-E)^2}{E}$	18	1	2.667	2

$\frac{4}{10} * 500 \Rightarrow 200$

$\frac{3}{10} * 500 \Rightarrow 150$

$\frac{2}{10} * 500 \Rightarrow 100$

$\frac{1}{10} * 500 \Rightarrow 50$

$\Sigma\left(\frac{(O-E)^2}{E}\right) \Rightarrow 23.667$

$\boxed{n \Rightarrow 4}$

$\alpha \Rightarrow 5\% \Rightarrow 0.05$

$df \Rightarrow n-1$

$\boxed{df \Rightarrow 3}$

$\chi^2_\alpha \Rightarrow 7.815$

$\chi^2 \Rightarrow 23.667$

$\chi^2 > \chi^2_\alpha$ H_0 is rejected.