* Fisher's Z - Transformation :-

To test the significance of an observed sample correlation coefficient from an uncorrelated bivariate normal population, t- is used. But in random sample of size n from a bivariate normal population in which $\rho \neq 0$, prof. R. A. Fisher proved that the distribution of "r" is by no means ~~normal~~ normal and in the neighbourhood of $\rho = \pm 1$, it's probability curve is extremely skewed even for large n. If $\rho \neq 0$, Fisher suggested the following distribution.

$$Z = \frac{1}{2} \log_e \left(\frac{1+r}{1-r} \right) = \tanh^{-1} r$$

and proved that even for small samples, the distribution Z is approximately normal with mean

$$\xi = \frac{1}{2} \log_e \left(\frac{1+\rho}{1-\rho} \right) = \tanh^{-1} \rho$$

and variance $\frac{1}{(n-3)}$ and for large values of n, say $n > 50$, the approximation is fairly good.

If $(Z - \xi)\sqrt{(n-3)} > 1.96$, Ho is rejected at 5% level of significance and if it is greater than 2.58, Ho is rejected at 1% level of significance.

Here, under null hypothesis H0,

$$U = \frac{Z - \xi}{1/\sqrt{n-3}} \quad \text{and} \quad S.E.(Z) = \frac{1}{\sqrt{n-3}}$$

* <u>Note:-</u> $\frac{1}{2}\log_e(a) = 1.1513 \log_{10}(a)$, a is real.

Que] A correlation coefficient of 0.72 is obtained from a sample of 29 pairs of observations. Can the sample be regarded as drawn from a bivariate normal population in which true correlation coefficient is 0.89.

<u>Solution:-</u>

Let, define hypothesis as

Null hypothesis (H0): There is no significant difference between $r = 0.72$ and $\rho = 0.80$ i.e. the sample can be regarded as drawn from the bivariate normal population with $\rho = 0.8$. Here $n = 29$ and

$$Z = \frac{1}{2}\log_e\left(\frac{1+r}{1-r}\right) = 1.1513\log_{10}\left(\frac{1+r}{1-r}\right)$$

$$\therefore Z = 1.1513\log_{10}\left(\frac{1+0.72}{1-0.72}\right) = 1.1513\log_{10}\left(\frac{1.72}{0.28}\right)$$

$$= 1.1513\log_{10}(6.14) = 1.1513 \times 0.7881$$

$$\boxed{Z = 0.9073}$$

$$\xi = \frac{1}{2}\log_e\left(\frac{1+\rho}{1-\rho}\right) = 1.1513\log_{10}\left(\frac{1+\rho}{1-\rho}\right)$$

$$\therefore \xi = 1.1513\log_{10}\left(\frac{1+0.80}{1-0.80}\right) = 1.1513\log_{10}\left(\frac{1.80}{0.20}\right)$$

$$= 1.1513 \cdot \log_{10}(9) = 1.1513 \times 0.9542$$

$$\therefore \boxed{\xi = 1.0986}$$

under H_0, the test statistic is

$$u = \frac{Z - \xi}{1/\sqrt{n-3}}$$

where $S.E.(Z) = \frac{1}{\sqrt{n-3}} = \frac{1}{\sqrt{29-3}} = \frac{1}{\sqrt{26}}$

$$= \frac{1}{5.0990} = 0.1961$$

$$\therefore S.E.(Z) = \frac{1}{\sqrt{n-3}} = 0.1961$$

$$\therefore u = \frac{Z - \xi}{1/\sqrt{n-3}} = \frac{0.9073 - 1.0986}{0.1961}$$

$$\therefore u = \frac{-0.1913}{0.1961} = -0.9755$$

$$\therefore |u| = |-0.9755| = 0.9755$$

Since $|u| < 1.96$, it is not significant at 5% level of significance and H_0 may be accepted. Hence the sample may be regarded as coming from a bivariate normal population with $\rho = 0.8$.

* Chi-square Test :- (χ^2) :-

χ^2 distribution has a large number of applications in statistics, some of which are enumerated below -

① To test if the hypothetical value of the population variance is $\sigma^2 = \sigma_0^2$ (say)

② To test the goodness of fit.

③ To test the independence of attributes.

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

⑤ To combine various probabilities obtained from independent experiments to give a single test of significance.

⑥ To test the homogeneity of independent estimates of the population correlation coefficient.

* Inferences About a population variance :-

Suppose we want to test if a random sample x_i ($i = 1, 2, -----, n$) has been drawn from a normal population with a specified variance $\sigma^2 = \sigma_0^2$ (say).

Under the null hypothesis (H_0) that the population variance is $\sigma^2 = \sigma_0^2$, the statistic

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{\sigma^2} \right)^2$$

Follows chi-square distribution with $(n-1)$ d.f.

By comparing the calculated value with the tabulated value of χ^2 for $(n-1)$ d.f. at certain level of significance (usually 5%) we may accept or reject the null hypothesis.

Que] It is believed that the precision (as measured by the variance) of an instrument is no more than 0.16. Write down the null and alternative hypothesis for testing this belief. Carry out the test at 1% level given 11 measurements of the same subject on the instrument.

2.5, 2.3, 2.4, 2.3, 2.5, 2.7, 2.5, 2.6, 2.6, 2.7, 2.5

Solution:-

Let, define the following hypothesis

Null Hypothesis (H0): $\sigma^2 = 0.16$

Alternative Hypothesis (H1): $\sigma^2 > 0.16$

By χ^2-square Test

$$\chi^2 = \frac{\Sigma(x-\bar{x})^2}{\sigma^2} \rightarrow (1)$$

Here, No. of observations $= n = 11$

$\therefore$ Degree of freedom (dF) $= n-1 = 11-1 = 10$

level of significance $= 1\% = 0.01$

$\therefore$ Tabulated χ^2 for 10 d.f. at 1% level of significance $= 23.209$

calculation of x^2 from given data –

X	$X-\overline{X}$	$(X-\overline{X})^2$
2.5	2.5 − 2.51 = −0.01	0.0001
2.3	2.3 − 2.51 = −0.21	0.0441
2.4	2.4 − 2.51 = −0.11	0.0121
2.3	2.3 − 2.51 = −0.21	0.0441
2.5	2.5 − 2.51 = −0.01	0.0001
2.7	2.7 − 2.51 = 0.19	0.0361
2.5	2.5 − 2.51 = −0.01	0.0001
2.6	2.6 − 2.51 = 0.09	0.0081
2.6	2.6 − 2.51 = 0.09	0.0081
2.7	2.7 − 2.51 = 0.19	0.0361
2.5	2.5 − 2.51 = −0.01	0.0001
$\Sigma x =$ 27.6		$\Sigma(x-\bar{x})^2 = 0.1891$

Here $n = 11$ and $\Sigma x = 27.6$

$\therefore \bar{x} = \frac{1}{n}\Sigma x = \frac{1}{11} \times 27.6 = 2.51$

$\therefore \boxed{\bar{x} = 2.51}$

under null hypothesis Ho, test statistic is

$$x^2 = \frac{\Sigma(x-\bar{x})^2}{\sigma^2} = \frac{0.1891}{0.16} = 1.182$$

$\therefore \boxed{x^2 = 1.182}$

$\because$ calculated $x^2 = 1.182 <$ tabulated $x^2 = 23.209$

$\therefore$ Ho is accepted and we conclude that the data are consistent with the hypothesis that the precision of the instrument is 0.6

Que] A manufacturer recorded the cut-off bias (volt) of a sample of 10 tubes as follows -

12.1, 12.3, 11.8, 12.0, 12.4, 12.0, 12.1, 11.9, 12.2, 12.2

The variability of cut-off bias for tubes of a standard type as measured by the standard deviation is 0.208 volts. Is the variability of the new tube with respect to cut-off bias less than that of the standard type? level of significance is 5%.

Solution! -

Here, $n=10$

standard deviation $(\sigma) = 0.208$

$\therefore$ variance $(\sigma^2) = (0.208)^2 =$

Now define the following hypothesis

Null hypothesis (H_0): $\sigma^2 = 0.0433$

Alternative hypothesis (H_1): $\sigma^2 < 0.0433$

Here d.f. $= n-1 = 10-1 = 9$

and level of significance $= 5\% = 0.05$

$\therefore$ From table, tabulated value of $\chi^2 = 16.919$

By χ^2-square Test,

under null hypothesis H_0,

$$\chi^2 = \frac{\sum_{i=1}^{n} (x-\bar{x})^2}{\sigma^2}$$

x	$x-\bar{x}$	$(x-\bar{x})^2$
12.1	12.1 − 12.1 = 0	0
12.3	12.3 − 12.1 = 0.2	0.04
11.8	11.8 − 12.1 = −0.3	0.09
12.0	12.0 − 12.1 = −0.1	0.01
12.4	12.4 − 12.1 = 0.3	0.09
12.0	12.0 − 12.1 = −0.1	0.01
12.1	12.1 − 12.1 = 0	0
11.9	11.9 − 12.1 = −0.2	0.04
12.2	12.2 − 12.1 = 0.1	0.01
12.2	12.2 − 12.1 = 0.1	0.01
$\Sigma x = 121$		$\Sigma(x-\bar{x})^2 = 0.3$

Here $n=10$ and $\Sigma x = 121$

$$\therefore \bar{x} = \frac{1}{n}\Sigma x = \frac{1}{10} \times 121 = 12.1$$

$$\therefore \boxed{\bar{x} = 12.1}$$

Here $\Sigma(x-\bar{x})^2 = 0.3$

$$\therefore \chi^2 = \frac{\sum_{i=1}^{10}(x-\bar{x})^2}{\sigma^2} = \frac{0.3}{0.0433} = 6.93$$

$$\therefore \boxed{\chi^2 = 6.93}$$

$\therefore$ calculated $\chi^2 = 6.93 <$ tabulated $\chi^2 = 16.919$

$\therefore$ We accept null hypothesis H_0 and we conclude that, $\sigma^2 = 0.0433$ i.e. the variability of the new tube with respect to cut-off-bias is not less than that of standard type.

15·6·2. Goodness of Fit Test. A very powerful test for testing the significance of the discrepancy between theory and experiment was given by Prof. Karl Pearson in 1900 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.

If f_i $(i = 1, 2, \ldots, n)$ is a set of observed (experimental) frequencies and e_i $(i = 1, 2, \ldots, n)$ is the corresponding set of expected (theoretical or hypothetical) frequencies, then Karl Pearson's chi-square, given by :

$$\chi^2 = \sum_{i=1}^{n} \left[\frac{(f_i - e_i)^2}{e_i} \right], \quad \left(\sum_{i=1}^{n} f_i = \sum_{i=1}^{n} e_i \right) \qquad \ldots(15\cdot15)$$

follows chi-square distribution with $(n-1)$ *d.f.*

Remark. This is an approximate test for large values of n. Conditions for the validity of the χ^2-test of goodness of fit have already been given in § 15·4 Remark 2 on page 15·12.

The goodness of fit test uses the chi-square distribution to determine if a hypothesized probability distribution for a population provides a good fit. Acceptance or rejection of the hypothesized population distribution is based upon differences between observed frequencies (f_i's) in a sample and the expected frequencies (e_i's) obtained under null hypothesis H_0.

Decision rule :Accept H_0 if $\chi^2 \le \chi^2_\alpha (n-1)$ and reject H_0 if $\chi^2 > \chi^2_\alpha (n-1)$, where χ^2 is the calculated value of chi-square obtained on using (15·15) and $\chi^2_\alpha (n-1)$ is the tabulated value of chi-square for $(n-1)$ *d.f.* and level of significance α.

Example 15·11. *The demand for a particular spare part in a factory was found to vary from day-to-day. In a sample study the following information was obtained :*

Days	:	*Mon.*	*Tues.*	*Wed.*	*Thurs.*	*Fri.*	*Sat.*
No. of parts demanded	:	1124	1125	1110	1120	1126	1115

Test the hypothesis that the number of parts demanded does not depend on the day of the week. (Given : the values of chi-square significance at 5, 6, 7, d.f. are respectively 11·07 12·59, 14·07 at the 5% level of significance.)

Solution. Here we set up the *null hypothesis*, H_0 that the number of parts demanded does not depend on the day of week.

Under the null hypothesis, the expected frequencies of the spare part demanded on each of the six days would be :

$$\frac{1}{6}(1124 + 1125 + 1110 + 1120 + 1126 + 1115) = \frac{6720}{6} = 1120$$

600÷4 = 24

TABLE 15·2 : CALCULATIONS FOR χ^2

Days	*Frequency.*		$(f_i - e_i)^2$	$\frac{(f_i - e_i)^2}{e_i}$
	Observed (f_i)	*Expected* (e_i)		
Mon.	1124	1120	16	0·014
Tues.	1125	1120	25	0·022
Wed.	1110	1120	100	0·089
Thurs.	1120	1120	0	0
Fri.	1126	1120	36	0·032
Sat.	1115	1120	25	0·022
Total	6720	6720		0·179

$$\chi^2 = \sum \frac{(f_i - e_i)^2}{e_i} = 0{\cdot}179$$

The number of degrees of freedom = 6 − 1 = 5 (since we are given 6 frequencies subjected to only one linear constraint : $\sum f_i = \sum e_i = 6720$)

The tabulated $\chi^2_{0{\cdot}05}$ for 5 *d.f.* = 11·07.

Since calculated value of χ^2 is less than the tabulated value, it is not significant and the null hypothesis may accepted at 5% level of significance. Hence we conclude that the number of parts demanded are same over the 6-day period.

Example 15·12. *The following figures show the distribution of digits in numbers chosen at random from a telephone directory :*

Digits :	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.

Solution. Here we set up the *null hypothesis* that the digits occur equally frequently in the directory.

Under the null hypothesis, the expected frequency for each of the digits 0, 1, 2, ..., 9 is 10,000/10 = 1000. The value of χ^2 is computed as follows :

TABLE 15·3 : CALCULATIONS FOR χ^2

Digits	Frequency		$(f_i-e_i)^2$	$\frac{(f_i-e_i)^2}{e_i}$
	Observed (f_i)	Expected (e_i)		
0	1026	1000	676	0·676
1	1107	1000	11449	11·449
2	997	1000	9	0·009
3	966	1000	1156	1·156
4	1075	1000	5625	5·625
5	933	1000	4489	4·489
6	1107	1000	11149	11·449
7	972	1000	784	0·784
8	964	1000	1296	1·296
9	853	1000	21609	21·609
Total	10,000	10,000		58·542

$$\chi^2 = \sum \frac{(f_i-e_i)^2}{e_i}$$
$$= 58{\cdot}542$$

The number of degrees of freedom

= Number of observations – Number of independent constraints

$= 10 - 1 = 9$

Tabulated $\chi^2_{0{\cdot}05}$ for 9 *d.f.* = 16·919

Since the calculated value of χ^2 is much greater than the tabulated value, it is highly significant and we reject the null hypothesis. Thus we conclude that the digits are not uniformly distributed in the directory.

Example 15·13. *A sample analysis of examination results of 200 MBA's was made. It was found that 46 students had failed, 68 secured a third division, 62 secured a second division and the rest were placed in first division. Are these figures commensurate with the general examination result which is in the ratio of 4 : 3 : 2 : 1 for various categories respectively ?*

Solution. Set up the *null hypothesis* that the observed figures do not differ significantly from the hypothetical frequencies which are in the ratio of 4 : 3 : 2 : 1. In other words *the given data are commensurate with the general examination result* which is in the ratio of 4 : 3 : 2 : 1 for the various categories.

Under the null hypothesis, the expected frequencies can be computed as shown in the adjoining table :

Category	Frequency	
	Observed (f_i)	Expected (e_i)
Failed	46	$\frac{4}{10} \times 200 = 80$
III Division	68	$\frac{3}{10} \times 200 = 60$
II Division	62	$\frac{2}{10} \times 200 = 40$
I Division	24	$\frac{1}{10} \times 200 = 20$
Total	200	200

TABLE 15·4 : CALCULATIONS FOR χ^2

Category	*Frequency* Observed (f_i)	Expected (e_i)	$(f_i - e_i)^2$	$\frac{(f_i - e_i)^2}{e_i}$
Failed	46	80	1156	14·450
III Division	68	60	64	1·067
II Division	62	40	484	12·100
I Division	24	20	16	0·800
Total	200	200		28·417

$$\chi^2 = \sum \frac{(f_i - e_i)^2}{e_i} = 28·417$$

$d.f. = 4 - 1 = 3$, tabulated $\chi^2_{0·05}$ for 3 $d.f. = 7·815$

Since the calculated value of χ^2 is greater than the tabulated value, it is significant and the null hypothesis is rejected at 5% level of significance. Hence we may conclude that data are not commensurate with the general examination result.

15·6·3. Test of Independence of Attributes—Contingency Tables. Let us consider two attributes A and B, A divided into r classes $A_1, A_2, \ldots, A_r$ and B divided into s classes $B_1, B_2, \ldots, B_s$. Such a classification in which attributes are divided into more than two classes is known as *manifold classification.* The various cell frequencies can be expressed in the following table known as $r \times s$ *manifold contingency table* where (A_i) is the number of persons possessing the attribute A_i, $(i = 1, 2, \ldots, r)$, (B_j) is the number of persons possessing the attribute B_j $(j = 1, 2, \ldots, s)$ and (A_iB_j) is the number of persons possessing both the attributes A_i and B_j, $(i = 1, 2, \ldots, r; j = 1, 2, \ldots, s)$.

Also $\sum\limits_{i=1}^{r} (A_i) = \sum\limits_{j=1}^{s} (B_j) = N$, where N is the total frequency.

TABLE 15·7 : $r \times s$ CONTINGENCY TABLE

B \ A	A_1	A_2	$\ldots$	A_i	$\ldots$	A_r	*Total*
B_1	(A_1B_1)	(A_2B_1)	$\ldots$	(A_iB_1)	$\ldots$	(A_rB_1)	(B_1)
B_2	(A_1B_2)	(A_2B_2)	$\ldots$	(A_iB_2)	$\ldots$	(A_rB_2)	(B_2)
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
B_j	(A_1B_j)	(A_2B_j)	$\ldots$	(A_iB_j)	$\ldots$	(A_rB_j)	(B_j)
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
B_s	(A_1B_s)	(A_2B_s)	$\ldots$	(A_iB_s)	$\ldots$	(A_rB_s)	(B_s)
Total	(A_1)	(A_2)	$\ldots$	(A_i)	$\ldots$	(A_r)	N

Example 15·16. *Two sample polls of votes for two candidates A and B for a public office are taken, one from among the residents of rural areas. The results are given in the adjoining table. Examine whether the nature of the area is related to voting preference in this election.*

Area	*Votes for*		*Total*
	A	*B*	
Rural	*620*	*380*	*1000*
Urban	*550*	*450*	*1000*
Total	*1170*	*830*	*2000*

Solution. Under the *null hypothesis* that the nature of the area is independent of the voting preference in the election, we get the expected frequencies as follows :

$$E(620) = \frac{1170 \times 1000}{2000} = 585, \qquad E(380) = \frac{830 \times 1000}{2000} = 415,$$

$$E(550) = \frac{1170 \times 1000}{2000} = 585, \quad \text{and} \quad E(450) = \frac{830 \times 1000}{2000} = 415$$

Aliter. In a 2×2 contingency table, since $d.f. = (2-1)(2-1) = 1$, only one of the cell frequencies can be filled up independently and the remaining will follow immediately, since the observed and theoretical marginal totals are fixed. Thus having obtained any one of the theoretical frequencies (say) $E(620) = 585$, the remaining theoretical frequencies can be easily obtained as follows :

$$E(380) = 1000 - 585 = 415, \; E(550) = 1170 - 585 = 585, \text{ and } E(450) = 1000 - 585 = 415.$$

$$\therefore \quad \chi^2 = \sum_i \left[\frac{(f_i - e_i)^2}{e_i} \right] = \frac{(620-585)^2}{585} + \frac{(380-415)^2}{415} + \frac{(550-585)^2}{585} + \frac{(450-415)^2}{415}$$

$$= (35)^2\left(\frac{1}{585}+\frac{1}{415}+\frac{1}{585}+\frac{1}{415}\right) = (1225)[2\times 0{\cdot}002409 + 2\times 0{\cdot}001709] = 10{\cdot}0891$$

Tabulated $\chi^2_{0.05}$ for $(2-1)(2-1) = 1$ *d.f.* is 3·841. Since calculated χ^2 is much greater than the tabulated value, it is highly significant and null hypothesis is rejected at 5% level of significance. Thus we conclude that nature of area is related to voting preference in the election.

Example 15·17. (2 × 2 CONTINGENCY TABLE). *For the 2 × 2 table,*

a	b
c	d

, prove that chi-square test of independence gives

$$\chi^2 = \frac{N(ad-bc)^2}{(a+c)\,(b+d)\,(a+b)\,(c+d)},\ N = a+b+c+d. \qquad \ldots(15{\cdot}18)$$

Solution. Under the hypothesis of independence of attributes,

$$E(a) = \frac{(a+b)\,(a+c)}{N}$$

$$E(b) = \frac{(a+b)\,(b+d)}{N}$$

$$E(c) = \frac{(a+c)\,(c+d)}{N}$$

and
$$E(d) = \frac{(b+d)\,(c+d)}{N}$$

a	b	$a+b$
c	d	$c+d$
$a+c$	$b+d$	N

$$\therefore \quad \chi^2 = \frac{[a-E(a)]^2}{E(a)} + \frac{[b-E(b)]^2}{E(b)} + \frac{[c-E(c)]^2}{E(c)} + \frac{[d-E(d)]^2}{E(d)} \qquad \ldots(*)$$

$$a - E(a) = a - \frac{(a+b)\,(a+c)}{N} = \frac{a(a+b+c+d)-(a^2+ac+ab+bc)}{N} = \frac{ad-bc}{N}$$

Similarly, we will get : $\quad b-E(b) = -\dfrac{ad-bc}{N} = c - E(c)\,; \quad d - E(d) = \dfrac{ad-bc}{N}$

Substituting in (*), we get

$$\chi^2 = \frac{(ad-bc)^2}{N^2}\left[\frac{1}{E(a)}+\frac{1}{E(b)}+\frac{1}{E(c)}+\frac{1}{E(d)}\right]$$

$$= \frac{(ad-bc)^2}{N}\left[\left\{\frac{1}{(a+b)(a+c)}+\frac{1}{(a+b)(b+d)}\right\}+\left\{\frac{1}{(a+c)\,(c+d)}+\frac{1}{(b+d)\,(c+d)}\right\}\right]$$

$$= \frac{(ad-bc)^2}{N}\left[\frac{b+d+a+c}{(a+b)(a+c)(b+d)}+\frac{b+d+a+c}{(a+c)\,(c+d)\,(b+d)}\right]$$

$$= (ad-bc)^2\left[\frac{c+d+a+b}{(a+b)(a+c)(b+d)(c+d)}\right] = \frac{N(ad-bc)^2}{(a+b)(a+c)(b+d)(c+d)}.$$

Remark. We can calculate the value of χ^2 for 2 × 2 contingency table by using (15·18) directly. The reader is advised to obtain the value of χ^2 in *Example 15·16* by using (15·18).

Example 15·18. *Out of 8,000 graduates in a town 800 are females, out of 1,600 graduate employees 120 are females. Use χ^2 to determine if any distinction is made in appointment on the basis of sex. Value of χ^2 at 5% level for one degree of freedom is 3·84.*

Solution. We set up the *Null hypothesis* that no distinction is made in appointment on the basis of sex, and test it against the *Alternative hypothesis* that distinction is made in appointment on the basis of sex.

The observed and expected frequencies are shown in the following table :

TABLE NO. OBSERVED FREQUENCIES | EXPECTED FREQUENCIES

	Employed	*Not employed*	*Total*	*Employed*	*Not employed*	*Total*
Male	1480	5720	7200	$\frac{7200 \times 1600}{8000}$ = 1440	7200 – 1440 = 5760	7200
Female	120	680	800	1600 – 1440 = 160	6400 – 5760 = 640	800
Total	1600	6400	8000	1600	6400	8000

TABLE 15·8 : CALCULATIONS FOR χ^2

Class	*Frequency* Observed (f_i)	*Frequency* Expected (e_i)	$(f_i - e_i)$	$\frac{(f_i - e_i)^2}{e_i}$
Male employed	1480	1440	40	$\frac{1600}{1440} = 1{\cdot}11$
Male unemployed	5720	5760	– 40	$\frac{1600}{5760} = 0{\cdot}28$
Female employed	120	160	– 40	$\frac{1600}{160} = 10{\cdot}00$
Female unemployed	680	640	40	$\frac{1600}{640} = 2{\cdot}50$

$$\chi^2 = \Sigma \frac{(f_i - e_i)^2}{e_i} = 13{\cdot}89$$

$$d.f. = (2-1)(2-2) = 1$$

Tabulated $\chi^2_{0.05}$ for 1 $d.f.$ = 3·841.

Conclusion. Since the calculated value of χ^2 (13·89) is much greater than the tabulated value of χ^2 (3·841), the value of χ^2 is highly significant and null hypothesis is rejected. Hence we conclude that distinction is made in appointment on the basis of sex.

TABLE 15.

SINGNIFICANT VALUES $\chi^2(\alpha)$ OF CHI-SQUARE DISTRIBUTION (RIGHT TAIL AREAS) FOR GIVEN PROBABILITY α,

where

$$P = P_r[\chi^2 > \chi_\upsilon^2(\alpha)] = \alpha$$

AND υ IS DEGREES OF FREEDOM (*d.f.*)

* χ^2-DISTRIBUTION VALUES OF $\chi_\upsilon^2(\alpha)$

Degrees of freedom (υ)	*Probability* (α)							
	0·995	0·99	0·995	0·95	0·05	0·025	0·01	0·005
1	0·000	0·000	0·001	0·004	3·841	5·024	6·635	7·879
2	0·010	0·020	0·051	0·103	5·991	7·378	9·210	10·597
3	0·072	0·115	0·216	0·352	7·815	9·348	11·345	12·838
4	0·207	0·297	0·484	0·711	9·488	11·143	13·277	14·860
5	0·412	0·554	0·831	1·145	11·070	12·832	15·086	16·750
6	0·676	0·872	1·237	1·634	12·592	14·449	16·812	18·548
7	0·989	1·239	1·690	2·167	14·067	16·013	18·475	20·278
8	1·344	1·646	2·180	2·733	15·507	17·535	20·090	21·955
9	1·735	2·088	2·700	3·325	16·919	19·023	21·666	23·589
10	2·156	2·558	3·247	3·940	18·307	20·483	23·209	25·188
11	2·603	3·053	3·816	4·575	19·675	21·920	24·725	26·757
12	3·074	3·571	4·404	5·226	21·026	23·337	26·217	28·300
13	3·565	4·107	5·009	5·892	22·362	24·736	24·888	29·819
14	4·075	4·660	5·629	6·571	23·685	26·119	29·141	31·319
15	4·601	5·229	6·262	7·261	24·996	27·488	30·578	32·801
16	5·142	5·812	6·908	7·962	26·296	28·845	32·000	34·267
17	5·697	6·408	7·564	8·672	27·587	30·191	33·409	35·718
18	6·265	7·015	8·231	9·390	28·869	31·526	34·805	37·156
19	6·844	7·633	8·907	10·117	30·144	32·852	36·191	38·582
20	7·434	8·260	9·591	10·851	31·410	34·170	37·566	39·997
21	8·034	8·897	10·283	11·591	32·671	35·479	38·932	41·401
22	8·643	9·542	10·982	12·338	33·924	36·781	40·289	42·796
23	9·260	10·196	11·688	13·091	35·172	38·076	41·638	44·181
24	9·886	10·856	12·401	13·848	36·415	39·364	42·980	45·558
25	10·520	11·524	13·120	14·611	37·652	40·646	44·314	46·928
26	11·160	12·198	13·844	15·379	38·885	41·923	45·642	48·290
27	11·808	12·879	14·573	16·151	40·113	43·194	46·963	49·645
28	12·461	13·565	15·308	16·928	41·337	44·461	48·278	50·993
29	13·121	14·256	16·047	17·708	42·557	45·722	49·588	52·336
30	13·787	14·953	16·791	18·493	43·773	46·979	50·892	53·672
40	20·706	22·164	24·433	26·509	55·759	59·342	63·691	66·766
50	27·991	29·707	32·357	34·764	67·505	71·420	76·154	79·490
60	35·535	37·485	40·482	43·188	79·082	83·298	88·379	91·952
70	43·275	45·442	48·758	51·739	90·531	95·023	100·425	104·215
80	51·172	53·540	57·153	60·391	101·879	106·629	112·329	116·321
90	59·196	61·754	65·647	69·126	113·145	118·136	124·116	128·299
100	67·328	70·065	74·222	77·929	124·342	129·561	135·807	140·169

For larger values of ν, quantity $\sqrt{2\chi^2} - \sqrt{2\nu - 1}$ may be used as a standard normal variable.

* ***Abridged from Table 8 of Biometrika Tables for Statisticians, Vol. I.***