$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{-6{\cdot}9}{\sqrt{\left(\frac{9{\cdot}6 + 27{\cdot}5}{17}\right)\left(\frac{1}{8} + \frac{1}{11}\right)}}$$

$$= -\frac{6{\cdot}9}{0{\cdot}6864} = -10.05$$

and $\nu = n_1 + n_2 - 2 = 17$

From the t-table $t_{0.05}$ $(\nu = 17) = 2.11$.

If H_0: $\bar{x}_1 = \bar{x}_2$ and H_1: $\bar{x}_1 \neq \bar{x}_2$, then H_0 is rejected, since $|t| > t_{0.05}$.

That is, the means of two samples (and so the populations) differ significantly.

Therefore, the two samples could not have been drawn from the same normal population.

Example 15 *The nicotine contents in two random samples of tobacco are given below.*

Sample 1 :	*21*	*24*	*25*	*26*	*27*	
Sample 2 :	*22*	*27*	*28*	*30*	*31*	*36*

Can you say that the two samples came from the same population?

Solution $\bar{x}_1$ = Mean of sample 1 = $\frac{123}{5} = 24.6$

$\bar{x}_2$ = Mean of sample 2 = $\frac{174}{6} = 29.0$

s_1^2 = Variance of sample 1 = $\frac{1}{5}\,\Sigma(x_i - 24.6)^2 = 4.24$

s_2^2 = Variance of sample 2 = $\frac{1}{6}\,\Sigma(x_i - 29.0)^2 = 18.0$

$\hat{\sigma}_1^2 = \frac{5}{4} \times 4.24 = 5.30$ and $\nu = 4$; $\hat{\sigma}_2^2 = \frac{6}{5} \times 18.0 = 21.60$ and $\nu = 5$

H_0: $\hat{\sigma}_1^2 = \hat{\sigma}_2^2$; H_1: $\hat{\sigma}_1^2 \neq \hat{\sigma}_2^2$

$$F = \frac{\hat{\sigma}_2^2}{\hat{\sigma}_1^2} = \frac{21{\cdot}60}{5{\cdot}30} = 4.07$$

$F_{0.05}\,(5, 4) = 6.26.$

Since $F < F_{0.05}$, H_0 is accepted. Therefore, the variances of the two populations can be regarded as equal.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{-4{\cdot}4}{\sqrt{\left(\frac{21{\cdot}2 + 108{\cdot}0}{9}\right)\left(\frac{1}{5} + \frac{1}{6}\right)}}$$

$$= \frac{-4 \cdot 4}{2 \cdot 2943} = -1.92$$

and $\nu = 9$.

From t-table, $F_{0.05}$ ($\nu = 9$) = 2.26.

If H_0: $\bar{x}_1 = \bar{x}_2$ and H_1: $\bar{x}_1 \neq \bar{x}_2$, H_0 is accepted, since $|t| < F_{0.05}$.

That is, the means of two samples (and hence the populations) do not differ significantly. Therefore, the two samples could have been drawn from the same normal population.

Exercise 8(B)

Part A (Short answer questions)

1. Write down the probability density of student's t-distribution.
2. State the important properties of the t-distribution.
3. Give any two uses of t-distribution.
4. What do you mean by degrees of freedom?
5. How will you get the critical value of t for a single-tailed test at LOS α?
6. What is the test statistic used to test the significance of the difference between small sample mean and population mean?
7. Give the 95% confidence interval of the population mean in terms of the mean and SD of a small sample.
8. What is the test statistic used to test the significance of the difference between the means of two small samples?
9. Give an estimate of the population variance in terms of variances of two small samples. What is the associated number of degrees of freedom?
10. What is the test statistic used to test the significance of the difference between the means of two small samples of the same size? What is the associated number of degree of freedom?
11. What is the test statistic used to test the significance of the difference between the means of two small samples of the same size, when the sample items are correlated?
12. Write down the probability density function of the F-distribution.
13. State the important properties of the F-distribution.
14. What is the use of F-distribution?
15. Why is the F-distribution associated with two numbers of degrees of freedom?

Part B

16. A random sample of 10 boys had the following IQs: 70, 120, 110, 101, 88, 83, 95, 98, 107, 100. Does the data support the assumption of a population

mean IQ of 100? Find a reasonable range in which most of the mean IQ values of samples of ten boys lie.

17. A random sample of 16 values from a normal population showed a mean of 103.75 cm and the sum of the squares of deviations from this mean is equal to 843.75 cm^2. Show that the assumption of a mean of 108.75 cm for the population is not reasonable. Obtain 95% and 99% fiducial limits for the same.
18. The mean weekly sales of soap bars in departmental stores is 145 bars store. After an advertising campaign, the mean weekly sales in 17 stores for a typical week increased to 155 and showed an SD of 16. Was the advertising campaign successful?
19. The annual rainfall at a certain place is normally distributed with mean 30. If the rainfalls during the past 8 years are 31.1, 30.7, 24.3, 28.1, 27.9, 32.2, 25.4 and 29.1, can we conclude that average rainfall during the past 8 years is less than the normal rainfall?
20. A machine is expected to produce nails of length 7 cm. A random sample of 10 nails was found to measure the following lengths: 7.2, 7.3, 7.1, 6.9, 6.8, 6.5, 6.9, 6.8, 7.1 and 7.2 cm. On the basis of this sample, what can be said about the reliability of the machine?
21. A random sample of 8 envelopes is taken from the letter box of a post office and their weights in grams are found to be 12.2, 11.9, 12.5, 12.3, 11.6, 11.7, 12.2 and 12.4. Find the 95% and 99% confidence limits for the mean weight of the envelopes in the letter box.
22. The average production of 16 workers in a factory was 107 with an SD of 9, while 12 workers in another comparable factory had an average production of 111 with an SD of 10. Can we say that the production rate of workers in the latter factory is more than that in the former factory?
23. Two different types of drugs A and B were tried on certain patients for increasing weight. Five persons were given drug A and 7 persons were given drug B. The increase in weight (in kg.) is given below:

Drug A :	3.6	5.5	5.9	4.1	1.4		
Drug B :	4.5	3.6	5.5	6.8	2.7	3.6	5.0

Do the two drugs differ significantly with regard to their effect in increasing weight?

24. Samples of 12 foremen in one division and 10 foremen in another division of a factory were selected at random and the following data were obtained:

	Division 1	*Division 2*
Sample size	12	10
Average monthly salary of foremen (Rs)	5250	4900
SD of salary (Rs)	152	165

Can you conclude that foremen in Division 1 get more salary than foremen in division 2?

25. Two independent groups of 10 children were tested to find how many digits they could repeat from memory after hearing them. The results are as follows:

Group A:	8	6	5	7	6	8	7	4	5	6
Group B:	10	6	7	8	6	9	7	6	7	7

Is the difference between mean scores of the two groups significant?

26. The following data represent the marks obtained by 12 students in 2 tests, one held before coaching and the other after coaching:

Test 1:	55	60	65	75	49	25	18	30	35	54	61	72
Test 2:	63	70	70	81	54	29	21	38	32	50	70	80

Does the data indicate that the coaching was effective in improving the performance of students?

27. In one sample of 8 items, the sum of the squares of deviations of the sample values from the sample mean was 84.4, and in another sample of 10 observations it was 102.6. Test whether this difference is significant at 5% level.

28. Two random samples drawn from two normal populations gave the following observations.

Sample 1:	20	16	26	27	23	22	18	24	25	19		
Sample 2:	17	23	32	25	22	24	28	18	31	33	20	27

Test whether the two populations have the same variance.

29. Two random samples gave the following results:

Sample no.	*Size*	*Mean*	*Variance*
1	16	440	40
2	25	460	42

Test whether the samples have been drawn from the same normal population.

30. Two random samples gave the following results:

Sample no.	*Size*	*Sum of the values*	*Sum of the squares of values*
1	10	150	2340
2	12	168	2460

Test whether the samples have been drawn from the same normal population.

Chi-Square Distribution

If $X_1, X_2, \ldots, X_n$ are normally distributed independent random variables, then it is known that $(X_1^2 + X_2^2 + \ldots + X_n^2)$ follows a probability distribution, called *chi-*

square (χ^2-distribution) distribution with n degrees of freedom. The probability density function of the χ^2-distribution is given by

$$f(\chi^2) = \frac{1}{2^{\nu/2}\sqrt{\left(\frac{\nu}{2}\right)}} \cdot \left(\chi^2\right)^{\nu/2-1} e^{-\chi^2/2} 0 < \chi^2 < \infty,$$

where ν is the number of degrees of freedom.

Properties of x^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. 8.4.
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.

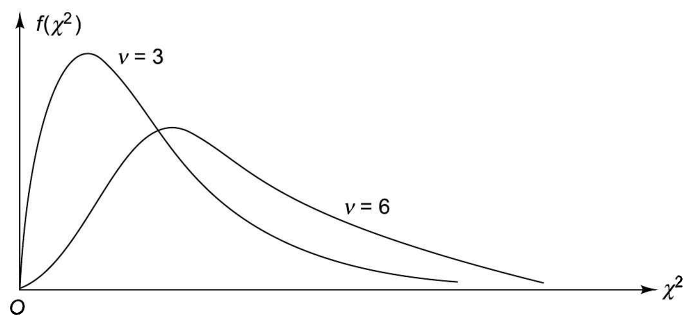

Fig. 8.4

4. As ν 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. That is, 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.

χ^2-Test of Goodness of Fit

On the basis of the hypothesis assumed about the population, we find the expected frequencies $E_i (i = 1, 2, \ldots, n)$, corresponding to the observed frequencies O_i $(i = 1, 2, \ldots, n)$ such that $\sum E_i = \sum O_i$. It is known that

$$\chi^2 = \sum_{i=1}^{n} \frac{\left(O_i - E_i\right)^2}{E_i}$$

follows approximately a χ^2-distribution with degrees of freedom equal to the number of independent frequencies. 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 can we 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.

The critical value of χ^2 for ν degrees of freedom at α LOS, denoted by $\chi^2_\nu(\alpha)$, is given by

$$P\,[\chi^2 > \chi^2_\nu(\alpha)] = \alpha$$

Critical values of the χ^2-distribution corresponding to a few important LOS and a range of values of ν are available in the form of a table called χ^2-table, which is given at the end of the chapter.

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. $4 \leq n \leq 16$.

χ^2-Test of Independence of Attributes

If the population is known to have two major attributes A and B, then A can be divided into m categories $A_1, A_2, \ldots, A_m$ and B can be divided into n categories $B_1, B_2, \ldots, B_n$. Accordingly the members of the population, and hence those of the sample, can be divided into mn classes. In this case, the sample data may be presented in the form of a matrix containing m rows and n columns and hence mn cells, and showing the observed frequencies O_{ij}, in the various cells, where $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. O_{ij} means the number of observed frequencies possessing the attributes A_i and B_j. The matrix or tabular form of the sample data, called an $(m \times n)$ *contingency table* is given below:

A			B				Row Total
	B_1	B_2	-	B_j	-	B_n	
A_1	O_{11}	O_{12}	-	O_{1j}	-	O_{1n}	O_{1*}
A_2	O_{21}	O_{22}	-	O_{2j}	-	O_{2n}	O_{2*}
:	-	-	-	-	-	-	-
A_i	O_{i1}	O_{i2}	-	O_{ij}	-	O_{in}	O_{i*}
:	-	-	-	-	-	-	-
A_m	O_{m1}	O_{m2}		O_{mj}	-	O_{mn}	O_{m*}
Column Total	O_{*1}	O_{*2}	-	O_{*j}	-	O_{*n}	N

Now, based on the null hypothesis H_0, i.e. the assumption that the two attributes A and B are independent, we compute the expected frequencies E_{ij} for various cells, using the following formula:

$$E_{ij} = \frac{O_{i*} \times O_{*j}}{N}, \; i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, n$$

i.e.
$$E_{ij} = \left(\frac{\text{Total of observed frequencies in the } ith \text{ row } \times \text{ total of a observed frequencies in the } j\text{th column}}{\text{Total of all cell frequencies}} \right)$$

Then we compute

$$\chi^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} \left\{ \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}} \right\}$$

The number of degrees of freedom for this χ^2 computed from the $(m \times n)$ contingency table is $\nu = (m - 1)(n - 1)$.

If $\chi^2 < \chi^2_\nu(\alpha)$, H_0 is accepted at α % LOS, i.e. the attributes A and B are independent.

If $\chi^2 > \chi^2_\nu(\alpha)$, H_0 is rejected at α % LOS, i.e. A and B are not independent.

Worked Examples 8(C)

Example 1 *The following table shows the distribution of digits in the numbers chosen at random from a telephone directory:*

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.

Solution H_0: The digits occur equally frequently, i.e. they follow a uniform distribution.

Based on H_0, we compute the expected frequencies.

The total number of digits = 10,000.

If the digits occur uniformly, then each digit will occur $\frac{10,000}{10} = 1000$ times.

O_i: 1026, 1107, ..., 853

E_i: 1000, 1000, ..., 1000

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

$$= \frac{1}{1000} [(26)^2 + (107)^2 + (-3)^2 + (-34)^2 + (75)^2 + (-67)^2 + (107)^2 + (-28)^2 + (-36)^2 + (-147)^2]$$

$$= 58.542$$

Since ΣE_i was taken equal to ΣO_i (i.e. an information from the sample), $v = n - 1 = 10 - 1 = 9$. From the χ^2–table,

$\chi^2_{0.05}$ $(n = 9) = 16.919$.

Since $\chi^2 > \chi^2_{0.05}$, H_0 is rejected. That is, the digits do not occur uniformly in the directory.

Example 2 *The following data give the number of air-craft accidents that occurred during the various days of a week.*

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

Test whether the accidents are uniformly distributed over the week.

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 = $\frac{90}{6} = 15$.

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

From the χ^2-table, $\chi^2_{0.05}$ $(v = 5) = 11.07$.

Since $\chi^2 < \chi^2_{0.05}$, H_0 is accepted. That is, accidents may be regarded to occur uniformly over the week.

Example 3 *The following data show defective articles produced by 4 machines:*

Machine:	*A*	*B*	*C*	*D*
Production time:	*1*	*1*	*2*	*3*
No. of detectives:	*12*	*30*	*63*	*98*

Do the figures indicate a significant difference in the performance of the machines?

Solution H_0: Production rates of the machines are the same.

Total number of detectives = 203.

Based on H_0, the expected numbers of defectives produced by the machines are

$$E_i: \quad \frac{1}{7} \times 203 \quad \frac{1}{7} \times 203 \quad \frac{2}{7} \times 203 \quad \frac{3}{7} \times 203$$

i.e.

$$E_i: \quad 29 \quad 29 \quad 58 \quad 87$$
$$O_i: \quad 12 \quad 30 \quad 63 \quad 98$$

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = \frac{17^2}{29} + \frac{1^2}{29} + \frac{5^2}{58} + \frac{11^2}{87} = 11.82$$

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

From the χ^2-table, $\chi^2_{0.05}$ $(v = 3) = 7.815$.

Since $\chi^2 > \chi^2_{0.05}$, H_0 is rejected. That is, there is significant difference in the performance of machines.

Example 4 *The following data represent the monthly sales (in Rs) of a certain retail store in a leap year. Examine if there is any seasonality in the sales.*
6100, 5600, 6350, 6050, 6250, 6200, 6300, 6250, 5800, 6000, 6150 and 6150.

Solution H_0: There is no seasonability in the sales, i.e. the daily sales are uniform throughout the year or the daily sales follow a uniform distribution.

Based on H_0, we compute the expected frequencies.

The total sales in the year = Rs 73,200.

If the daily sales are uniform, then the sales on each day $= \dfrac{73{,}200}{366}$ = Rs 200.

O_i: 6100, 5600, 6350, 6050, 6250, 6200, 6300, 6250, 5800, 6000, 6150, 6150.

Assuming that the months are taken in the usual calendar order, namely, January, February, etc., the expected monthly sales are:

E_i: 6200, 5800, 6200, 6000, 6200, 6000, 6200, 6200, 6000, 6200, 6000, 6200

Then
$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

$$= \frac{(-100)^2}{6200} + \frac{(-200)^2}{5800} + \ldots + \frac{(-50)^2}{6200} = 38.913$$

Since $\sum E_i$ was found as $\sum O_i$ from the sample, $\nu = n - 1 = 12 - 1 = 11$.

From the χ^2-table $\chi^2_{0.05}$ $(\nu = 11) = 19.675$.

Since $\chi^2 > \chi^2_{0.05}$, H_0 is rejected. That is, the daily sales are not uniform throughout the year.

Example 5 *Theory predicts that the proportion of beans in 4 groups A, B, C, D should be 9 : 3 : 3 : 1. In an experiment among 1600 beans, the numbers in the 4 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 4 groups are in the ratio 9 : 3 : 3 : 1.

Based on H_0, the expected numbers of beans in the 4 groups are as follows:

	E_i:	$\frac{9}{16} \times 1600$	$\frac{3}{16} \times 1600$	$\frac{3}{16} \times 1600$	$\frac{1}{16} \times 1600$
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 $\sum E_i = \sum O_i$, $\nu = 4 - 1 = 3$.

From the χ^2-table, $\chi^2_{0.05}(\nu = 3) = 7.82$.

Since $\chi^2 < \chi^2_{0.05}$, H_0 is accepted. That is, the experimental data support the theory.

Example 6 *A survey of 320 families with 5 children revealed the following distribution:*

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

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

Solution H_0: Male and female births are equally probable, i.e. P (male birth) = $p = 1/2$ and P (female birth) = $q = 1/2$.

Based on H_0, the probability that a family of 5 children has r male children

$= 5C_r \left(\frac{1}{2}\right)^5$ (by binomial law).

$\therefore$ Expected number of families having r male children $= 320 \times 5C_r \times \frac{1}{2^5}$

$= 10 \times 5C_r$.

Thus	E_i:	10	50	100	100	50	10
	O_i:	12	40	88	110	56	14

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = \frac{2^2}{10} + \frac{10^2}{50} + \frac{12^2}{100} + \frac{10^2}{100} + \frac{6^2}{50} + \frac{4^2}{10}$$

$$= 7.16$$

We have used the sample data to get $\sum E_i$ only. The values of p and q have not been found by using the sample data.

$\therefore \quad v = n - 1 = 6 - 1 = 5$ and $\chi^2_{0.05}(v = 5) = 11.07$

Since $\chi^2 < \chi^2_{0.05}$, H_0 is accepted. That is, male and female births are equally probable.

Example 7 *Twelve dice were thrown 4096 times and a throw of 6 was considered a success. The observed frequencies were as given below:*

No. of successes:	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7 and over*
Frequency:	*447*	*1145*	*1180*	*796*	*380*	*115*	*25*	*8*

Test whether the dice were unbiased.

Solution H_0: All the dice were unbiased,

i.e. P (getting 6) $= p = \frac{1}{6}$. Therefore, $q = \frac{5}{6}$.

Based on H_0, the probability of getting exactly r successes

$$= 12C_r\, p^r q^{12-r} \quad (r = 0, 1, 2, \ldots, 12)$$

Therefore, expected number of times in which r successes are obtained:

$$E_r = 4096 \times 12C_r \left(\frac{1}{6}\right)^r \cdot \left(\frac{5}{6}\right)^{12-r}$$

$$= 4096 \times 12C_r \times \frac{5^{12-r}}{6^{12}} \quad (r = 0, 1, 2, \ldots, 12)$$

i.e.

$$E_0 = N\,(0 \text{ success}) = N\,(r = 0) = 459.39$$
$$E_1 = N\,(r = 1) = 1102.54$$
$$E_2 = N\,(r = 2) = 1212.80$$
$$E_3 = N\,(r = 3) = 808.53$$
$$E_4 = N\,(r = 4) = 363.84$$
$$E_5 = N\,(r = 5) = 116.43$$
$$E_6 = N\,(r = 6) = 27.17$$
$$E_7 = N\,(r \geq 7) = 5.30$$

Converting E_i's into whole numbers subject to the condition that $\sum E_i = 4096$, we get

E_i:	459	1103	1213	809	364	116	27	5
O_i:	497	1145	1180	796	380	115	25	8

Since E and O corresponding to the last class i.e. 5 and 8 are less than 10, we combine the last 2 classes and consider as a single class.

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = \frac{12^2}{459} + \frac{42^2}{1103} + \frac{33^2}{1213} + \frac{13^2}{809} + \frac{16^2}{364} + \frac{1^2}{116} + \frac{1^2}{32}$$

$= 3.76$

$v = n - 1$ (since only $\sum E_i$ has been found using the sample data).

$= 7 - 1$ (n must be taken as the number of classes after combination of end classes, if any)

$= 6$

and $\chi^2_{0.05}$ ($v = 6$) = 12.59, from the χ^2-table. Since $\chi^2 < \chi^2_{0.05}$, H_0 is accepted, i.e. the dice were unbiased.

Example 8 *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

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

Solution 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

The first class is combined with the second and the last 2 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 $n = n - k = 4 - 2 = 2$ and $\chi^2_{0.05} (\nu = 2) = 5.99$, from the χ^2-table.

Since $\chi^2 > \chi^2_{0.05}$, 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.

Example 9 *Fit a Poisson distribution for the following distribution and also test the goodness of fit.*

x:	0	1	2	3	4	5	Total
f:	142	156	69	27	5	1	400

Solution *To find the Poisson distribution whose probability law is*

$$P(X = r) = \frac{e^{-\lambda} \lambda^r}{r!}, \; r = 0, 1, 2, \ldots$$

we require λ, which is the mean of the Poisson distribution.

We find the mean of the given distribution and assume it as λ.

x:	0	1	2	3	4	5	Total
f:	142	156	69	27	5	1	400
fx:	0	156	138	81	20	5	400

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{400}{400} = 1 = \lambda$$

The expected frequencies are given by

$$\frac{N \cdot e^{-\lambda} \lambda^r}{r!} \text{ or } \frac{400 \times e^{-1}}{r!}, \; r = 0, 1, 2, \ldots, \infty$$

Thus

E_i:	147.15	147.15	73.58	24.53	6.13	1.23

The values of E_i are very small for $i = 6, 7, \ldots$ and hence neglected.

Converting the values of E_i's into whole numbers such that $\sum E_i = 400$, we get

E_i:	147	147	74	25	6	1

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

O_i:	142	156	69	27	5	1

The last 3 classes are combined into one, so that the expected frequency in that class may be ≥ 10. Thus, after regrouping, we have

O_i:	142	156	69	33
E_i:	147	147	74	32

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = \frac{5^2}{147} + \frac{9^2}{147} + \frac{5^2}{74} + \frac{1^2}{32}$$

$$= 1.09$$

We have used the sample data to find $\sum E_i$ and λ. Hence, $n = n - k = 4 - 2 = 2$

From the χ^2-table, $\chi^2_{0.05}$ $(\nu = 2) = 5.99$.

Since $\chi^2 < \chi^2_{00.5}$, H_0, which assumes that the given distribution is nearly Poisson, is accepted. That is, the Poisson fit for the given distribution is satisfactory.

Example 10 *Test the normality of the following distribution by using χ^2-test of goodness of fit*

x:	*125*	*135*	*145*	*155*	*165*	*175*	*185*	*195*	*205*	*Total*
f:	*1*	*1*	*14*	*22*	*25*	*19*	*13*	*3*	*2*	*100*

Solution Let us first fit a normal distribution to the given data and then test the goodness of fit.

To fit a normal distribution and hence find the expected frequencies, we require the density function of the normal distribution which involves the mean and SD. Let us now compute the mean $\bar{x}$ and SD s of the sample distribution and assume them as μ and σ.

x	f	$d = \frac{x - 165}{10}$	fd	fd^2
125	1	– 4	– 4	16
135	1	– 3	– 3	9
145	14	– 2	– 28	56
155	22	– 1	– 22	22
165	25	0	0	0
175	19	1	19	19
185	13	2	26	52
195	3	3	9	27
205	2	4	8	32
Total	100	–	5	233

$$\bar{x} = A + \frac{c}{N} \sum f d = 165 + \frac{10}{100} \times 5 = 165.5$$

$$s^2 = c^2 \left\{ \frac{1}{N} \sum f d^2 - \left(\frac{1}{N} \sum f d \right)^2 \right\} = 10^2\,(2.33 - 0.0025)$$

$$= 232.75$$

$\therefore \qquad s = 15.26$

Therefore, the density function of the normal distribution which fits the given distribution is $f(x) = \dfrac{1}{15 \cdot 26\sqrt{2\pi}}\, e^{-(x-165\cdot 5)^2/465\cdot 5}$.

To find the expected frequency corresponding to a given x, we find $y = f(x)$ and multiply y by the class-width and then by the total frequency N.

Note $y = \dfrac{1}{\sigma\sqrt{2\pi}}\, e^{-(x-\mu)^2/2\sigma^2}$. *If we put* $\dfrac{x-\mu}{\sigma} = z$,

then $y = \dfrac{1}{\sigma}\left\{\dfrac{1}{\sqrt{2\pi}}\, e^{-z^2/2}\right\} = \dfrac{\phi(z)}{\sigma}$, *where $\phi(z)$ is density function of the standard normal distribution. Values of $\phi(z)$ are obtained from the normal table given at the end of this chapter.*

x	$y = \dfrac{x - 165\cdot 5}{15\cdot 26}$	$\phi(z)$	$\dfrac{c\cdot\phi(z)}{\sigma} = \dfrac{10\phi(z)}{15\cdot 26}$	*Expected frequency* $= N\phi(z)/\sigma$
125	– 2.65	0.0119	0.0078	0.78
135	– 2.00	0.0540	0.0354	3.54
145	– 1.34	0.1626	0.1066	10.66
155	– 0.69	0.3144	0.2060	20.60
165	– 0.03	0.3988	0.2613	26.13
175	0.62	0.3292	0.2157	21.57
185	1.28	0.1758	0.1152	11.52
195	1.93	0.0620	0.0406	4.06
205	2.59	0.0139	0.0091	0.91

Converting the expected frequencies as whole numbers such that $\sum E_i = 100$, we get

E_i: 1 3 11 21 26 22 11 4 1

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

Combining the end classes so as to make the individual frequencies greater than 10,

E_i:	15	21	26	22	16
O_i:	16	22	25	19	18

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = \frac{1^2}{15} + \frac{1^2}{21} + \frac{1^2}{26} + \frac{3^2}{22} + \frac{2^2}{16}$$

$$= 0.82$$

We have used the sample data to find $\sum E_i$, μ and σ.

Hence $v = n - k = 5 - 3 = 2$.

From the χ^2-table, $\chi^2_{0.05}$ $(v = 2) = 5.99$.

Since $\chi^2 < \chi^2_{0.05}$, H_0, which assumes that the given distribution is nearly normal, is accepted. That is, the normal fit for the given distribution is satisfactory.

Example 11 *The following data is collected on two characters. Based on this, can you say that there is no relation between smoking and literacy?*

	Smokers	*Non-smokers*
Literates:	*83*	*57*
Illiterates:	*45*	*68*

Solution H_0: Literacy and smoking habit are independent

	Smokers	*Non-smokers*	*Total*
Literates	83	57	140
Illiterates	45	68	113
Total	128	125	253

O	*E*	*E (rounded)*	$(O-E)^2/E$
83	$\frac{128 \times 140}{253} = 70.83$	71	$12^2/71 = 2.03$
57	$\frac{125 \times 140}{253} = 69.17$	69	$12^2/69 = 2.09$
45	$\frac{128 \times 113}{253} = 57.17$	57	$12^2/57 = 2.53$
68	$\frac{125 \times 113}{253} = 55.83$	56	$12^2/56 = 2.57$
			$\chi^2 = 9.22$

$$\begin{aligned} v &= (m-1)(n-1) \\ &= (2 - 1 (2 - 1) = 1. \end{aligned}$$

From the χ^2-table, $\chi^2_{0.05}$ $(v = 1) = 3.84$

Since $\chi^2 > \chi^2_{0.05}$, H_0 is rejected. That is, there is some association between literacy and smoking.

Example 12 *Prove that the value of χ^2 for the 2×2 contingency table* $\begin{array}{|c|c|} \hline a & b \\ \hline c & d \\ \hline \end{array}$ *is given by*

$$\chi^2 = \frac{N(ad - bc)^2}{(a+b)(c+d)(a+c)(b+d)}, \text{ where } N = a + b + c + d.$$

Hence compute χ^2 for the 2×2 contingency table given in Example 11.

Solution The value of E corresponding to the cell in which $O = a$ is given by

$$E = \frac{(a+b)(a+c)}{(a+b+c+d)}.$$

Therefore, the value of χ^2 corresponding to this cell is given by

$$\chi^2 = \left\{ a - \frac{(a+b)(a+c)}{a+b+c+d} \right\}^2 \div \frac{(a+b)(a+c)}{(a+b+c+d)}$$

$$= \frac{\{a(a+b+c+d) - (a+b)(a+c)\}^2}{N(a+b)(a+c)}$$

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

Similarly the values of χ^2 are found out for the other three cells.

$\therefore$ χ^2 for the table

$$= \frac{(ad-bc)^2}{N}$$

$$\left[\frac{1}{(a+b)(a+c)} + \frac{1}{(a+b)(b+d)} + \frac{1}{(a+c)(c+d)} + \frac{1}{(b+d)(c+d)} \right]$$

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

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

$$= \frac{N(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)} \tag{1}$$

Using (1) for the contingency table in Example 11, we get

$$\chi^2 = \frac{253(83 \times 68 - 45 \times 57)^2}{140 \times 113 \times 128 \times 125} = 9.48$$

Example 13 *Two batches each of 12 animals are taken for test of inoculation. One batch was inoculated and the other batch was not inoculated. The numbers of dead and surviving animals are given in the following table for both cases. Can the inoculation be regarded as effective against the disease? Make Yate's correction for continuity of χ^2.*

	Dead	*Survived*	*Total*
Inoculated	2	10	12
Not inoculated	8	4	12
Total	10	14	24

Note on Yate's correction for continuity of χ^2.

Solution The χ^2-table was prepared using the theoretical χ^2-distribution which is continuous, whereas the approximate values of χ^2 that we are using are discrete. To rectify this defect, Yates has shown that, when

$$\chi^2 = \sum \left[\frac{\left\{ \left| 0_i - E_i \right| - \frac{1}{2} \right\}^2}{E_i} \right]$$

is used, the χ^2-approximation is improved. Yate's correction is used only when $v = 1$ and hence for a 2×2 contingency table. It is used only when some cell frequency is small, i.e. less than 5.

In the present problem, two cell frequencies are less than 5 each. Hence we apply Yate's correction:

O	E	$\lvert O - E \rvert - 0.5$	$\{\lvert O - E \rvert - 0.5\}^2/E$
2	$\frac{12 \times 10}{24} = 5$	2.5	6.25/5 = 1.25
10	$\frac{12 \times 14}{24} = 7$	2.5	6.25/7 = 0.89
8	$\frac{12 \times 10}{24} = 5$	2.5	6.25/5 = 1.25
4	$\frac{12 \times 14}{24} = 7$	2.5	6.25/7 = 0.89
	$v = (2 - 1)(2 - 1) = 1$		$\chi^2 = 4.28$

From the χ^2-table, $\chi^2_{0.05}$ $(v = 1) = 3.84$.

If H_0: Inoculation and effect on the diseases are independent, then H_0 is rejected as $\chi^2 > \chi^2_{0.05}$ i.e. Inoculation can be regarded as effective against the disease.

Note *Even if Yate's correction is not made, we would have arrived at the same conclusion.*

Example 14 *A total number of 3759 individuals were interviewed in a public opinion survey on a political proposal. Of them, 1872 were men and the rest women. A total of 2257 individuals were in favour of the proposal and 917 were opposed to it. A total of 243 men were undecided and 442 women were opposed* *to the proposal. Do you justify or contradict the hypothesis that there is no association between sex and attitude?*

Solution A careful analysis of the problem results in the following contingency table:

	Favoured	*Opposed*	*Undecided*	*Total*
Men	1154	475	243	1872
Women	1103	442	342	1887
Total	2257	917	585	3759

H_0: Sex and attitude are independent, i.e. there is no association between sex and attitude.

O	*E (rounded E)*	$(O-E)^2/E$
1154	$\dfrac{1872 \times 2257}{3759} \simeq 1124$	$30^2/1124 = 0.80$
475	$\dfrac{1872 \times 917}{3759} \simeq 457$	$18^2/457 = 0.71$
243	$\dfrac{1872 \times 585}{3759} \simeq 291$	$48^2/291 = 7.92$
1103	$\dfrac{1887 \times 2257}{3759} \simeq 1133$	$30^2/1133 = 0.79$
442	$\dfrac{1887 \times 917}{3759} \simeq 460$	$18^2/460 = 0.70$
342	$\dfrac{1887 \times 585}{3759} \simeq 294$	$48^2/294 = 7.84$
	$\nu = (3-1)(2-1) = 2$	$\chi^2 = 18.76$

From the χ^2-table, $\chi^2_{0.05}$ $(\nu = 2) = 5.99$.

Since $\chi^2 > \chi^2_{0.05}$, H_0 is rejected. Therefroe, sex and attitude are not independent, i.e. there is some association between sex and attitude.

Example 15 *The following table gives for a sample of married women, the level of education and the marriage adjustment score:*

Level of	*Marriage adjustment*				
Education	*Very low*	*Low*	*High*	*Very high*	*Total*
College	*24*	*97*	*62*	*58*	*241*
High school	*22*	*28*	*30*	*41*	*121*
Middle school	*32*	*10*	*11*	*20*	*73*
Total	*78*	*135*	*103*	*119*	*435*

Can you conclude from the above data that the higher the level of education, the greater is the degree of adjustment in marriage?

Solution H_0: There is no relation between the level of education and adjustment in marriage.

$$v = (4 - 1)(3 - 1) = 6$$

$$\chi^2_{0.05}(v = 6) = 12.59$$

O	E (rounded)	$(O-E)^2/E$
24	43	$19^2/43 = 8.40$
97	75	$22^2/75 = 6.45$
62	57	$5^2/57 = 0.44$
58	66	$8^2/66 = 0.97$
22	22	$0^2/22 = 0.00$
28	37	$9^2/37 = 2.19$
30	29	$1^2/29 = 0.03$
41	33	$8^2/33 = 1.94$
32	13	$19^2/13 = 27.77$
10	23	$13^2/23 = 7.35$
11	17	$6^2/17 = 2.12$
20	20	$0^2/20 = 0.00$
	$\chi^2_{0.05}(v = 6) = 12.59$	$\chi^2 = 57.66$

Since $\chi^2 > \chi^2_{0.05}$, H_0 is rejected. That is, the level of education and adjustment in marriage are associated.

Thus we may conclude that the higher the level of education, the greater is the degree of adjustment in marriage.

Exercise 8(C)

Part A (Short answer questions)

1. Define chi-square distribution.
2. Write down the probability density function of the χ^2-distribution.
3. State the important properties of χ^2-distribution.
4. Give 2 uses of χ^2-distribution.
5. What is χ^2-test of goodness of fit?
6. State the conditions under which χ^2-test of goodness of fit is valid.
7. What is χ^2-test of independence of attributes?
8. What is contingency table?
9. Write down the value of χ^2 for a 2×2 contingency table with cell frequencies a, b, c and d.
10. What is Yate's correction for continuity of χ^2?

Part B

11. In 250 digits from the lottery numbers, the frequencies of the digits were as follows:

Digit:	0	1	2	3	4	5	6	7	8	9
Frequency:	23	25	20	23	23	22	29	25	33	27

Test the hypothesis that the digits were randomly drawn.

12. The following table gives the number of fatal road accidents that occurred during the 7 days of the week. Find whether the accidents are uniformly distributed over the week.

Day:	Sun	Mon	Tues	Wed	Thu	Fri	Sat
Number:	8	14	16	12	11	14	9

13. In 120 throws of a single die, the following distribution of faces are obtained:

Face:	1	2	3	4	5	6
Frequency:	30	25	18	10	22	15

Do these results support the equal probability hypothesis?

14. The number of demands for a particular spare part in a shop was found to vary from day to day. In a sample study, the following information was obtained:

Day:	Mon	Tues	Wed	Thu	Fri	Sat
No. of demands:	124	125	110	120	126	115

Test the hypothesis that the number of parts demanded does not depend on the day of the week.

15. According to genetic theory, children having one parent of blood type M and the other of blood type N will always be one of the three types-M, MN and N and the average proportions of these types will be 1 : 2 : 1. Out of 300 children, having one M parent and one N parent, 30 per cent were found to be of type M, 45 per cent of type MN and the remaining of type N. Test the genetic theory by χ^2-test.

16. Five coins are tossed 256 times. The number of heads observed is given below. Examine if the coins are true.

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

17. Five dice were thrown 243 times and the numbers of times 1 or 2 was thrown (x) are given below:

x:	0	1	2	3	4	5
Frequency:	30	75	76	47	13	2

Examine if the dice were unbiased.

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

x:	0	1	2	3	4
f:	5	29	36	25	5

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

x:	0	1	2	3	4	5	6	7	8	9
f:	3	8	11	15	16	14	12	11	9	1

20. Fit a Poisson distribution for the following distribution and also test the goodness of fit.

x:	0	1	2	3	4	5	6	7
f:	314	335	204	86	29	9	3	0

21. Fit a Poisson distribution for the following distribution and also test the goodness of fit.

x:	0	1	2	3	4
f:	123	59	14	3	1

22. The figures given below are (i) the observed frequencies of a distribution and (ii) the expected frequencies of the normal distribution having the same mean, SD and total frequency as in (i).
(i) 1, 12, 66, 220, 495, 792, 924, 792, 495, 220, 66, 12, 1
(ii) 2, 15, 66, 210, 484, 799, 943, 799, 484, 210, 66, 15, 2
Do you think that the normal distribution provides a good fit to the data?
23. Fit a normal distribution to the following data and find also the goodness of fit.

x:	4	6	8	10	12	14	16	18	20	22	24
f:	1	7	15	22	35	43	38	20	13	5	1

24. In an epidemic of certain disease, 92 children contacted the disease. Of these 41 received no treatment and of these 10 showed after-effects. Of the remainder who did receive the treatment, 17 showed after-effects. Test the hypothesis that the treatment was not effective.
25. Out of 1660 candidates who appeared for a competitive examination, 422 were successful. Out of these, 256 had attended a coaching class and 150 of them came out successful. Examine whether coaching was effective as regards the success in the examination.
26. In a pre-poll survey, out of 1000 rural voters, 620 favoured A and the rest B. Out of 1000 urban voters, 450 favoured B and the rest A. Examine whether the nature of the area is related to voting preference.
27. The following information was obtained in a sample of 40 small general shops:

	Shops in urban areas	*Shops in rural areas*
Owned by men	17	18
Owned by women	3	12

Can it be said that there are more women owners in rural areas than in urban areas? Use Yate's correction for continuity.
28. A certain drug is claimed to be effective in curing cold. In an experiment on 500 persons with cold, half of them were given the drug and half of them were given the sugar pills. The patients' reaction to the treatment are recorded in the following table:

	Helped	*Harmed*	*No effect*
Drug	150	30	70
Sugar pills	130	40	80

On the basis of this data, can it be concluded that the drug and sugar pills differ significantly in curing cold?

29. A survey of radio listeners' preference for two types of music under various age groups gave the following information.

Type of music	*Age group*		
	19-25	26-35	Above 36
Carnatic music	80	60	90
Film music	210	325	44
Indifferent	16	45	132

Is preference for type of music influenced by age?

30. The table given below shows the results of a survey in which 250 respondents were classified according to the levels of education and attitude towards students' agitation in a certain town. Test whether the two criteria of classification are independent.

	Attitude		
Education	*Against*	*Neutral*	*For*
Middle school:	40	25	5
High school:	40	20	5
College:	30	15	30
Postgraduate:	15	15	10

Test whether the two criteria of classification are independent.

Table of normal ordinates and areas:

z	*Ordinate*	*Area* $= \int_0^z$	z	*Ordinate*	*Area* $= \int_0^z$
0.00	0.3989	0.0000	1.55	0.1200	0.4394
0.05	0.3984	0.0199	1.60	0.1109	0.4452
0.10	0.3970	0.0398	1.65	0.1023	0.4505
0.15	0.3945	0.0596	1.70	0.941	0.4554
0.20	0.3910	0.0793	1.75	0.0863	0.4099
0.25	0.3867	0.0987	1.80	0.0790	0.4641
0.30	0.3814	0.1179	1.85	0.0721	0.4678
0.35	0.3752	0.1368	1.90	0.0656	0.4719
0.40	0.3683	0.1554	1.95	0.0596	0.4744
0.45	0.3605	0.1736	2.00	0.0540	0.4773
0.50	0.3521	0.1910	2.05	0.0488	0.4798
0.55	0.3429	0.2088	2.10	0.0440	0.4821
0.60	0.3332	0.2258	2.15	0.0396	0.4842
0.65	0.3230	0.2422	2.20	0.0355	0.4861

(*Contd*)

z	*Ordinate*	*Area* = $\int_0^z$	z	*Ordinate*	*Area* = $\int_0^z$
0.70	0.3123	0.2080	2.25	0.0317	0.4878
0.75	0.3011	0.2734	2.30	0.0283	0.4893
0.80	0.2897	0.2882	2.35	0.0252	0.4906
0.85	0.2780	0.3023	2.40	0.0224	0.4918
0.90	0.2661	0.3159	2.45	0.0224	0.4929
0.95	0.2541	0.3289	2.50	0.0198	0.4938
1.00	0.2420	0.3413	2.55	0.0175	0.4946
1.05	0.2299	0.3531	2.60	0.0155	0.4953
1.10	0.2179	0.3643	2.65	0.0136	0.4960
1.15	0.2059	0.3749	2.70	0.0119	0.4965
1.20	0.1942	0.3849	2.75	0.0104	0.4970
1.25	0.1827	0.3944	2.80	0.0091	0.4974
1.30	0.1714	0.4032	2.85	0.0079	0.4978
1.35	0.1604	0.4115	2.90	0.0060	0.4981
1.40	0.1497	0.4192	2.95	0.0051	0.4984
1.45	0.1394	0.4265	3.00	0.0044	0.4987
1.50	0.1295	0.4332	3.05	0.0038	0.4989

***t*-Table**

	Probability				
n	0.9	0.1	0.05	0.02	0.01
1	0.158	6.314	12.706	31.821	63.657
2	0.142	2.920	4.303	6.965	9.925
3	0.137	2.353	3.182	4.541	5.841
4	0.134	2.132	2.776	3.747	4.604
5	0.132	2.015	2.571	3.365	4.032
6	0.131	1.943	2.447	3.143	3.707
7	0.130	1.895	2.365	2.998	3.496
8	0.129	1.860	2.306	2.896	3.355
9	0.129	1.833	2.262	2.821	3.250
10	0.129	1.812	2.228	2.764	3.169
11	0.128	1.796	2.201	2.718	3.106
12	0.128	1.782	2.179	2.681	3.055
13	0.128	1.771	2.160	2.650	3.012
14	0.128	1.761	2.145	2.624	2.977
15	0.128	1.753	2.131	2.602	2.947
16	0.128	1.746	2.120	2.583	2.921
17	0.128	1.740	2.110	2.567	2.898
18	0.127	1.734	2.101	2.552	2.878
19	0.127	1.729	2.093	2.539	2.861
20	0.127	1.725	2.086	2.528	2.845
21	0.127	1.721	2.080	2.518	2.831
22	0.127	1.717	2.074	2.508	2.819
23	0.127	1.714	2.069	2.500	2.807
24	0.127	1.711	2.064	2.492	2.797

(*Contd*)

	Probability				
n	0.9	0.1	0.05	0.02	0.01
25	0.127	1.708	2.060	2.485	2.787
30	0.127	1.697	2.042	2.457	2.750
40	0.126	1.684	2.021	2.423	2.704
60	0.126	1.671	2.000	2.390	2.660
120	0.126	1.658	1.980	2.358	2.617
∞	0.126	1.645	1.960	2.326	2.576

χ^2-Table

	Probability					
n	0.99	0.95	0.10	0.05	0.02	0.01
1	0.000157	0.00393	2.706	3.841	5.412	6.635
2	0.0201	0.103	4.605	5.991	7.824	9.210
3	0.115	2.352	6.251	4.815	9.837	11.345
4	0.297	0.711	7.779	9.488	11.668	13.277
5	0.554	1.145	9.236	11.070	13.388	15.086
6	0.872	1.635	10.645	12.592	15.033	16.812
7	1.238	2.167	12.017	14.067	16.622	18.475
8	1.646	2.733	13.362	15.507	18.168	20.090
9	2.088	3.325	14.684	16.919	19.670	21.666
10	2.558	3.940	15.987	18.307	21.161	23.209
11	3.053	4.575	17.275	19.675	22.618	24.725
12	3.571	5.226	18.549	21.026	24.054	26.217
13	4.107	5.892	19.812	22.362	25.472	27.688
14	4.660	6.571	21.064	23.685	26.873	29.141
15	5.229	7.261	23.307	24.996	28.259	30.578
16	5.812	7.962	23.542	26.296	29.633	32.000
17	6.408	8.672	24.768	27.587	30.995	33.409
18	7.015	9.390	25.989	28.869	32.346	34.805
19	7.633	10.117	27.204	30.114	33.687	36.191
20	8.260	10.851	28.412	31.410	35.020	37.566
21	8.897	11.581	29.615	32.671	36.343	38.932
22	9.542	12.338	30.813	33.924	37.659	40.289
23	10.196	13.091	32.007	35.172	38.968	41.638
24	10.856	13.848	33.196	36.415	40.270	42.980
25	11.524	14.611	34.382	37.652	41.566	44.314
26	12.198	15.379	35.563	38.885	42.856	45.642
27	12.879	16.151	36.741	40.113	44.140	46.963
28	13.565	16.928	37.916	41.337	45.419	48.278
29	14.256	17.708	39.087	42.557	46.693	49.588
30	14.953	18.493	40.256	43.773	47.962	50.892

For larger values of n, the expression $\sqrt{2\chi^2} - \sqrt{2n-1}$ may be used as a normal variate with unit variance.

Table of F (Variance ratio)—1 Per Cent Points

v_2 \ v_1	1	2	3	4	5	6	8	12	24	∞
1	4052	4999	5403	5625	5764	5859	5981	6106	6234	6366
2	98.49	99.01	99.17	99.25	99.30	99.33	99.36	99.42	99.46	99.50
3	34.12	30.81	29.46	28.71	28.24	27.91	27.49	27.05	26.60	26.12
4	21.20	18.00	16.69	15.98	15.52	15.21	14.80	14.37	13.93	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.27	9.89	9.47	9.02
6	13.74	10.92	9.78	9.15	8.75	8.47	8.10	7.72	7.31	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.84	6.47	6.07	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.03	5.67	5.28	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.47	5.11	4.73	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.06	4.71	4.33	3.91
11	9.65	7.20	6.22	5.67	5.32	5.07	4.74	4.40	4.02	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.50	4.16	3.78	3.36
13	9.07	6.70	5.74	5.20	4.86	4.62	4.30	3.96	3.59	3.16
14	8.86	6.51	5.56	5.03	4.69	4.46	4.14	3.80	3.43	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.00	3.67	3.29	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	3.89	3.55	3.18	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.79	3.45	3.08	2.65
18	8.28	6.01	5.09	4.58	4.25	4.01	3.71	3.37	3.00	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.63	3.30	2.92	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.56	3.23	2.86	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.51	3.17	2.80	2.36
22	7.94	5.72	4.82	4.31	3.99	3.76	3.45	3.12	2.75	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.41	3.07	2.70	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.36	3.03	2.66	2.21
25	7.77	5.57	4.68	4.18	3.86	3.63	3.32	2.99	2.62	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.29	2.96	2.58	2.13
27	7.68	5.49	4.60	4.11	3.79	3.56	3.26	2.93	2.55	2.10
28	7.64	5.45	4.57	4.07	3.76	3.53	3.23	2.90	2.52	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.20	2.87	2.49	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.17	2.84	2.47	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	2.99	2.66	2.29	1.81
60	7.08	4.98	4.13	3.65	3.34	3.12	2.82	2.50	2.12	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.66	2.34	1.95	1.38
∞	6.64	4.60	3.78	3.32	3.02	2.80	2.51	2.18	1.79	1.00

Table of F—5 Per Cent Points

v_2 \ v_1	1	2	3	4	5	6	8	12	24	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	238.9	243.9	249.0	253.4
2	18.51	19.00	19.16	19.25	19.30	19.33	19.37	19.41	19.45	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.84	8.74	8.64	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.04	5.91	5.77	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.82	4.68	4.53	4.36

(*Contd*)

	v_1									
v_2	1	2	3	4	5	6	8	12	24	∞
6	5.99	5.14	4.76	4.53	4.39	4.28	4.15	4.00	3.84	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.73	3.57	3.41	3.23
8	5.32	4.46	4.07	3.87	3.69	3.58	3.44	3.28	3.12	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.23	3.07	2.90	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.07	2.91	2.74	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	2.95	2.79	2.61	2.40
12	4.75	3.88	3.49	3.26	3.11	3.00	2.85	2.69	2.50	2.30
13	4.67	3.80	3.41	3.18	3.02	2.92	2.77	2.60	2.42	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.70	2.53	2.35	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.64	2.48	2.29	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.59	2.42	2.24	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.55	2.38	2.19	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.51	2.34	2.15	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.48	2.31	2.11	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.45	2.28	2.08	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.42	2.25	2.05	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.40	2.23	2.03	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.38	2.20	2.00	1.76
24	4.26	3.40	3.10	2.78	2.62	2.51	2.36	2.18	1.98	1.73
25	4.24	3.38	2.99	2.76	2.60	2.49	2.34	2.16	1.96	1.71
26	4.22	3.37	2.98	2.74	2.59	2.47	2.32	2.15	1.95	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.30	2.13	1.93	1.67
28	4.20	3.34	2.95	2.71	2.56	2.44	2.29	2.12	1.91	1.65
29	4.18	3.33	2.93	2.70	2.54	2.43	2.28	2.10	1.90	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.27	2.09	1.89	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.18	2.00	1.79	1.51
60	4.00	3.15	2.76	2.52	2.37	2.25	2.10	1.92	1.70	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.02	1.83	1.61	1.25
∞	3.84	2.99	2.60	2.37	2.21	2.09	1.94	1.75	1.52	1.00

Answers

Exercise 8(A)

29. $z = 2.83$; significant
30. $z = 4.5$; the coin is not fair
31. $z = 1.79$; claim cannot be supported.
32. No, since z (= 1.40) < z_α (= 1.645).
34. No; (16.7, 27.0)
35. difference due to sampling fluctuations
36. $z = 0.725$; not significant
37. $z = 6.49$; Yes; Yes.
38. Yes, since z (= 1.80) > $z_{.05}$ (= 1.645)
39. $z = 2.56$; the difference cannot be hidden

40. z (= 0.104) < $z_{.05}$ (= 1.645); the machine has improved
41. z (= 3.17) > $z_{.05}$ (= 1.96); difference significant.
42. $|z|$ = 1.02; the claim is valid
43. z = 3.12; No
44. No, since z = 5 : 96.8
45. Claim cannot be true as z = 1.89 and $z_{0.05}$ = 1.645
46. Yes, since z (= 1.6) < $z_{0.05}$ (= 1.645)
47. 41
48. (57.2, 58.8) and (57.0, 59.0)
49. No, since z = 2.58
50. No, since z = 8.82
51. No, since z = 1.32
52. Yes, since z (= 4.78) > $z_{1\%}$ (= 2.33)
53. Yes, at 5% level, since $|z|$ (= 1.937) > $z_{0.05}$ (= 1.645); and No, at 1% level, since $|z|$ (= 1.937) < $|z_\alpha|$ (= 2.33)
54. (1.98, 6.02)
55. H_0: $\mu_1 - \mu_2$ = 35 accepted, as z = 1.90
57. Yes, as z = 1.71
58. Yes, as $|z|$ (= 2.5) > $z_{1\%}$ (= 2.33)
59. Yes, as z = 1.70
60. No, as z = 3.61

Exercise 8(B)

16. $|t|$ = 0.62; Yes; 83.66 < μ < 110.74
17. $|t|$ = 2.67; (99.76, 107.74) and (98.22, 109.29)
18. t = 2.5; the campaign was successful
19. $|t|$ = 1.44; $t_{0.1}$ = 1.90; $\bar{x}$ is not less than μ
20. $|t|$ = 0.26; machine is reliable
21. (11.82, 12.38); (11.69, 12.51)
22. t = – 1.067; No
23. $|t|$ = 0.424; No
24. t = 4.33; Yes
25. $|t|$ = 1.85; Not significant
26. t = 4.0; coaching was effective
27. F = 1.057; Not significant
28. The populations have the same variance
29. No, though the difference between variances is not significant, the difference between the mean is significant.
30. Yes, as the differences between the means and between the variance are not significant.

Exercise 8(C)

11. χ^2 = 5.2; ν = 9; digits randomly drawn

12. $\chi^2 = 4.17$; $v = 6$; accidents occur uniformly.
13. $\chi^2 = 12.9$; $v = 5$; equal probability hypothesis is refuted.
14. $\chi^2 = 1.68$; $v = 5$; the demand does not depend on the day of the week
15. $\chi^2 = 4.5$; $v = 2$; genetic theory may be correct
16. $\chi^2 = 3.54$; $v = 3$; coins are true
17. $\chi^2 = 2.76$; $v = 4$; dice are unbiased
18. E_i: 7, 26, 37, 24, 6; $\chi^2 = 0.06$; $v = 1$; binomial fit is good. (E_i: 1, 5, 11, 18, 21, 19, 13, 8, 3, 1)
19. $\chi^2 = 11.30$; $v = 4$; binomial fit is not satisfactory;
20. E_i: 301, 362, 217, 87, 26, 6, 1; $\chi^2 = 5.40$; $v = 4$; Poisson fit is good
21. E_i: 121, 61, 15, 3, 0 ; $\chi^2 = 0.99$; $v = 1$; Poisson fit is good
22. $x^2 = 3.84$; $v = 8$; Normal fit is good
23. E_i: 2, 5, 13, 25, 37, 42, 36, 23, 12, 4, 1; $\chi^2 = 1.68$; $v = 4$; Normal fit is good
24. $\chi^2 = 0.85$; $v = 1$; no association between treatment and after-effect
25. $\chi^2 = 176.12$; $v = 1$; coaching was effective
26. $\chi^2 = 10.09$; $v = 1$; some relation between area and voting preference
27. $\chi^2 = 2.48$; $v = 1$; no, as there is no relation between area and sex of ownership
28. $\chi^2 = 3.52$; $v = 2$; do not differ significantly
29. $\chi^2 = 373.40$; $v = 4$; preference for type of music influenced by age
30. $\chi^2 = 35.42$; $v = 6$; the two criteria are not independent