F-test

A random variable X is said to follow F distribution, if its probability density function is given by

$$f(F) = \frac{(v_1/v_2)^{v_1/2}}{\beta\left(\frac{v_1}{2}, \frac{v_2}{2}\right)} \cdot \frac{F^{\frac{v_1}{2}-1}}{\left(1 + \frac{v_1 F}{v_2}\right)^{(v_1+v_2)/2}}, \quad F > 0,$$

where v_1 and v_2 are the degrees of freedom of samples.

Use of F distribution:

F distribution is used to test the equality of the variance of the populations from which two small samples have been drawn.

F test of significance of the difference between population variances

To test the significance of the difference between population variances, we shall first find their estimates, $\widehat{\sigma_1}^2$ and $\widehat{\sigma_2}^2$ based on the sample variances s_1^2 and s_2^2. We compute estimates by the following formulas:

$$\widehat{\sigma_1}^2 = \frac{n_1 s_1^2}{n_1 - 1},$$

$$\widehat{\sigma_2}^2 = \frac{n_2 s_2^2}{n_2 - 1}.$$

Test statistics is $F = \frac{\widehat{\sigma_2}^2}{\widehat{\sigma_1}^2}$, where $\widehat{\sigma_1}^2 < \widehat{\sigma_2}^2$ or $F = \frac{\widehat{\sigma_1}^2}{\widehat{\sigma_2}^2}$, where $\widehat{\sigma_2}^2 < \widehat{\sigma_1}^2$

The value of F is greater than 1. Then, compare calculated value F with the table value $F_{v_1, v_2}(\alpha)$ (in case of $\widehat{\sigma_2}^2 < \widehat{\sigma_1}^2$) or $F_{v_2, v_1}(\alpha)$ (in case of $\widehat{\sigma_1}^2 < \widehat{\sigma_2}^2$) by using F–table.

If the calculated value F is less than the table value, then fail to reject null hypothesis.

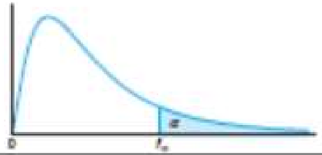

Table A.6 Critical Values of the F-Distribution

$f_{0.05}(v_1, v_2)$

v_2 \ v_1	1	2	3	4	5	6	7	8	9
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

Reproduced from Table 18 of *Biometrika Tables for Statisticians*, Vol. I, by permission of E.S. Pearson and the Biometrika Trustees.

Table A.6 (continued) Critical Values of the F-Distribution

$f_{0.05}(v_1, v_2)$

v_2 \ v_1	10	12	15	20	24	30	40	60	120	∞
1	241.88	243.91	245.95	248.01	249.05	250.10	251.14	252.20	253.25	254.31
2	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

Table A.6 (continued) Critical Values of the F-Distribution

	$f_{0.01}(v_1, v_2)$								
	v_1								
v_2	1	2	3	4	5	6	7	8	9
1	4052.18	4999.50	5403.35	5624.58	5763.65	5858.99	5928.36	5981.07	6022.47
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41

Table A.6 (continued) Critical Values of the F-Distribution

	$f_{0.01}(v_1, v_2)$									
	v_1									
v_2	10	12	15	20	24	30	40	60	120	∞
1	6055.85	6106.32	6157.28	6208.73	6234.63	6260.65	6286.78	6313.03	6339.39	6365.86
2	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
4	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

Example: Two independent samples of 7 and 6 items respectively had the following values of the variable.

Sample A: 28, 30, 32, 33, 33, 29, 34.

Sample B: 29, 30, 30, 24, 27, 29.

Do the two estimates of population variance differ significantly at 5% LOS?

Solution: We have $n_1 = 7$, $n_2 = 6$.

Degrees of freedom = $n_1 - 1, n_2 - 1 = 6, 5$.

Level of significance = 5%

Null hypothesis H_0 : The two horses have the same running capacity, i.e. $\sigma_1^2 = \sigma_2^2$

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

Test statistics is $F = \frac{\widehat{\sigma_2}^2}{\widehat{\sigma_1}^2} = \frac{5.368}{5.224} = 1.02$

Table value of F for (5, 6) degrees of freedom at 5% LS is 4.39.

Note: After computing two variance estimates $\widehat{\sigma_1}^2$ and $\widehat{\sigma_2}^2$ for two different samples our next job is to compute F-test statistic value. If you look at the F-test statistic formula and it's provided condition always in numerator we are taking highest value and in denominator lowest value of estimates.

Since the calculated value of F is less than the table value of F for d.fs (5, 6) at 5% L.S we accept H_0.

The two horses have the same running capacity.

Example: A sample of size 13 gave an estimated population variance of 3.0, while another sample of size 15 gave an estimate of 2.5. Could both samples be from populations with the same variance?

Solution Here, $n_1 = 13$, $\hat{\sigma}_1^2 = 3.0$ and $v_1 = 12$, $n_2 = 15$, $\hat{\sigma}_2^2 = 2.5$ and $v_2 = 14$.

H_0: $\hat{\sigma}_1^2 = \hat{\sigma}_2^2$, i.e. the two samples have been drawn from populations with the same variance. $H_1 : \hat{\sigma}_1^2 \neq \hat{\sigma}_2^2$.

Let LOS. be 5%.

$$F = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} = \frac{3 \cdot 0}{2 \cdot 5} = 1.2$$

$$v_1 = 12 \quad \text{and} \quad v_2 = 14.$$

$F_{0.05\%}$ ($v_1 = 12$, $v_2 = 14$) = 2.53, from the F-table. Since $F < F_{0.05}$, H_0 is accepted. That is the two samples could have come from two normal populations with the same variance.

χ^2-Test

In this study, we introduce chi-square distribution, the measure of which enables us to find the degree of discrepancy between the observed and expected frequencies is due to error of sampling or due to chance.

- The chi-square is denoted by the symbol χ^2. It is always positive. The value of chi-square lies between 0 and ∞.
- Since chi-square is not derived from the observation in a population, it is not a parameter. The chi-square test is not a parametric test.
- Chi-square is computed on the basis of frequencies in a sample and the value of chi-square so obtained is a statistic.

χ^2-Test is defined as

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

where O_i= Observed frequencies and E_i= Expected frequencies.

χ^2-Distribution

Let samples of size *n* be drawn from a normal population with standard deviation σ. If for each sample we calculate χ^2 a sampling distribution of χ^2 can be obtained. It is given by

$$f(\chi^2) = \frac{1}{2^{v/2}\sqrt{\frac{v}{2}}}(\chi^2)^{\frac{v}{2}-1}e^{-\frac{\chi^2}{2}},$$

where $0 < \chi^2 < \infty$ and *v* is the number of degrees of freedom.

Conditions for using χ^2-Test

- The total number of observations used in this test must be large.
- Each of the observations making up the sample for the χ^2 test should be independent of each other.
- The test is wholly dependent on the degrees of freedom.
- The frequencies used in a χ^2 test should be absolute and not relative in terms.
- The expected frequency of any item or cell should not be less than 5. If it is less than 5, then the frequencies from the adjacent items or cells should be pooled together in order to make it 5 or more than 5 (preferably not less than 10).
- The observations collected for χ^2 must be based on the method of random sampling.

Uses of χ^2-Test

χ^2-test is an important test. We require only the degrees of freedom for using this test. It is used

- as a test of goodness of fit,
- as a test of independence of attributes, and
- as a test of homogeneity. (This concept is not in your syllabus.)

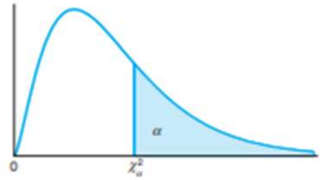

Table A.5 Critical Values of the Chi-Squared Distribution

v	α 0.995	0.99	0.98	0.975	0.95	0.90	0.80	0.75	0.70	0.50
1	0.0^4393	0.0^3157	0.0^3628	0.0^3982	0.00393	0.0158	0.0642	0.102	0.148	0.455
2	0.0100	0.0201	0.0404	0.0506	0.103	0.211	0.446	0.575	0.713	1.386
3	0.0717	0.115	0.185	0.216	0.352	0.584	1.005	1.213	1.424	2.366
4	0.207	0.297	0.429	0.484	0.711	1.064	1.649	1.923	2.195	3.357
5	0.412	0.554	0.752	0.831	1.145	1.610	2.343	2.675	3.000	4.351
6	0.676	0.872	1.134	1.237	1.635	2.204	3.070	3.455	3.828	5.348
7	0.989	1.239	1.564	1.690	2.167	2.833	3.822	4.255	4.671	6.346
8	1.344	1.647	2.032	2.180	2.733	3.490	4.594	5.071	5.527	7.344
9	1.735	2.088	2.532	2.700	3.325	4.168	5.380	5.899	6.393	8.343
10	2.156	2.558	3.059	3.247	3.940	4.865	6.179	6.737	7.267	9.342
11	2.603	3.053	3.609	3.816	4.575	5.578	6.989	7.584	8.148	10.341
12	3.074	3.571	4.178	4.404	5.226	6.304	7.807	8.438	9.034	11.340
13	3.565	4.107	4.765	5.009	5.892	7.041	8.634	9.299	9.926	12.340
14	4.075	4.660	5.368	5.629	6.571	7.790	9.467	10.165	10.821	13.339
15	4.601	5.229	5.985	6.262	7.261	8.547	10.307	11.037	11.721	14.339
16	5.142	5.812	6.614	6.908	7.962	9.312	11.152	11.912	12.624	15.338
17	5.697	6.408	7.255	7.564	8.672	10.085	12.002	12.792	13.531	16.338
18	6.265	7.015	7.906	8.231	9.390	10.865	12.857	13.675	14.440	17.338
19	6.844	7.633	8.567	8.907	10.117	11.651	13.716	14.562	15.352	18.338
20	7.434	8.260	9.237	9.591	10.851	12.443	14.578	15.452	16.266	19.337
21	8.034	8.897	9.915	10.283	11.591	13.240	15.445	16.344	17.182	20.337
22	8.643	9.542	10.600	10.982	12.338	14.041	16.314	17.240	18.101	21.337
23	9.260	10.196	11.293	11.689	13.091	14.848	17.187	18.137	19.021	22.337
24	9.886	10.856	11.992	12.401	13.848	15.659	18.062	19.037	19.943	23.337
25	10.520	11.524	12.697	13.120	14.611	16.473	18.940	19.939	20.867	24.337
26	11.160	12.198	13.409	13.844	15.379	17.292	19.820	20.843	21.792	25.336
27	11.808	12.878	14.125	14.573	16.151	18.114	20.703	21.749	22.719	26.336
28	12.461	13.565	14.847	15.308	16.928	18.939	21.588	22.657	23.647	27.336
29	13.121	14.256	15.574	16.047	17.708	19.768	22.475	23.567	24.577	28.336
30	13.787	14.953	16.306	16.791	18.493	20.599	23.364	24.478	25.508	29.336
40	20.707	22.164	23.838	24.433	26.509	29.051	32.345	33.66	34.872	39.335
50	27.991	29.707	31.664	32.357	34.764	37.689	41.449	42.942	44.313	49.335
60	35.534	37.485	39.699	40.482	43.188	46.459	50.641	52.294	53.809	59.335

Table A.5 (continued) Critical Values of the Chi-Squared Distribution

v	α 0.30	0.25	0.20	0.10	0.05	0.025	0.02	0.01	0.005	0.001
1	1.074	1.323	1.642	2.706	3.841	5.024	5.412	6.635	7.879	10.827
2	2.408	2.773	3.219	4.605	5.991	7.378	7.824	9.210	10.597	13.815
3	3.665	4.108	4.642	6.251	7.815	9.348	9.837	11.345	12.838	16.266
4	4.878	5.385	5.989	7.779	9.488	11.143	11.668	13.277	14.860	18.466
5	6.064	6.626	7.289	9.236	11.070	12.832	13.388	15.086	16.750	20.515
6	7.231	7.841	8.558	10.645	12.592	14.449	15.033	16.812	18.548	22.457
7	8.383	9.037	9.803	12.017	14.067	16.013	16.622	18.475	20.278	24.321
8	9.524	10.219	11.030	13.362	15.507	17.535	18.168	20.090	21.955	26.124
9	10.656	11.389	12.242	14.684	16.919	19.023	19.679	21.666	23.589	27.877
10	11.781	12.549	13.442	15.987	18.307	20.483	21.161	23.209	25.188	29.588
11	12.899	13.701	14.631	17.275	19.675	21.920	22.618	24.725	26.757	31.264
12	14.011	14.845	15.812	18.549	21.026	23.337	24.054	26.217	28.300	32.909
13	15.119	15.984	16.985	19.812	22.362	24.736	25.471	27.688	29.819	34.527
14	16.222	17.117	18.151	21.064	23.685	26.119	26.873	29.141	31.319	36.124
15	17.322	18.245	19.311	22.307	24.996	27.488	28.259	30.578	32.801	37.698
16	18.418	19.369	20.465	23.542	26.296	28.845	29.633	32.000	34.267	39.252
17	19.511	20.489	21.615	24.769	27.587	30.191	30.995	33.409	35.718	40.791
18	20.601	21.605	22.760	25.989	28.869	31.526	32.346	34.805	37.156	42.312
19	21.689	22.718	23.900	27.204	30.144	32.852	33.687	36.191	38.582	43.819
20	22.775	23.828	25.038	28.412	31.410	34.170	35.020	37.566	39.997	45.314
21	23.858	24.935	26.171	29.615	32.671	35.479	36.343	38.932	41.401	46.796
22	24.939	26.039	27.301	30.813	33.924	36.781	37.659	40.289	42.796	48.268
23	26.018	27.141	28.429	32.007	35.172	38.076	38.968	41.638	44.181	49.728
24	27.096	28.241	29.553	33.196	36.415	39.364	40.270	42.980	45.558	51.179
25	28.172	29.339	30.675	34.382	37.652	40.646	41.566	44.314	46.928	52.619
26	29.246	30.435	31.795	35.563	38.885	41.923	42.856	45.642	48.290	54.051
27	30.319	31.528	32.912	36.741	40.113	43.195	44.140	46.963	49.645	55.475
28	31.391	32.620	34.027	37.916	41.337	44.461	45.419	48.278	50.994	56.892
29	32.461	33.711	35.139	39.087	42.557	45.722	46.693	49.588	52.335	58.301
30	33.530	34.800	36.250	40.256	43.773	46.979	47.962	50.892	53.672	59.702
40	44.165	45.616	47.269	51.805	55.758	59.342	60.436	63.691	66.766	73.403
50	54.723	56.334	58.164	63.167	67.505	71.420	72.613	76.154	79.490	86.660
60	65.226	66.981	68.972	74.397	79.082	83.298	84.58	88.379	91.952	99.608

χ^2-test as a test of goodness of fit

χ^2-test is applied as a test of goodness of fit to determine whether the actual (i.e. observed) frequencies are close to the expected (i.e. theoretical) frequencies. The degrees of freedom in this case are $v = n - 1$, where n is the number of observations.

Problem 1: In 90 throws of a die, face 1 turned 9 times, face 2 or 3 turned 24 times, face 4 or 5 turned 36 times and 6 turned 18 times. Test at 10% level if the die is honest.

Solution: H_0 : The die is honest

H_1 : The die is not honest

Expected frequencies for each face= $90 * \frac{1}{6} = 15$

Level of significance= 10% = 0.1

Degree of freedom=4 − 1 = 3

χ^2-value for 3 degrees of freedom at 10% level of significance = 6.25.

χ^2-test as a test of goodness of fit (Conti...)

Face turned	Observed O_i	Expected E_i	$\frac{(O_i - E_i)^2}{E_i}$
1	9	15	2.4
2 or 3	27	30	0.3
4 or 5	36	30	1.2
6	18	15	0.6
Total	90	90	4.5

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

Since the calculated value of χ^2 is less than the table value at 10% level of significance and for 3 degrees of freedom, we accept the null hypothesis, and conclude that the die honest.

Problem 2: A sample analysis of examination results of 500 students was made. It was found that 230 students had failed. 160 had secured a third class, 80 were placed in second class and 30 got first class. Do these figures commensurate with the general examination results which is in the ratio 4 : 3 : 2 : 1 for various categories respectively?

Solution: H_0 : The observed results commensurate with the general examination results

H_1 : It is not true that the observed results commensurate with the general examination results.

Level of significance=5%. Degrees of freedom=4 − 1 = 3.

The total frequency=N = 500.

Table value χ^2 for 3 df at 5% level of significance = 7.81.

Dividing 500 in the ratio 4 : 3 : 2 : 1, we get 200, 150, 100 and 50. Therefore, the expected frequencies are 200, 150, 100 and 50 corresponding to the observed frequencies 230, 160, 80 and 30.

Class/division	Observed O_i	Expected E_i	$\frac{(O_i - E_i)^2}{E_i}$
Failed	230	200	4.500
Third	160	150	0.666
Second	80	100	4.000
First	30	50	8.000

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

Since the calculated value of χ^2 is greater than the table value of χ^2 , the null hypothesis is rejected.

χ^2-test as a test of goodness of fit (Conti...)

Problem 3:The table below gives the number of aircraft accidents that occurred during the various days of the week. Test whether the accidents are uniformly distributed over the week.

Days	Mon	Tue	Wed	Thu	Fri	Sat
No. of accidents	14	18	12	11	15	14

Problem 4:Fit a Poisson distribution to the following data and test the goodness of fit

x:	0	1	2	3	4	5	6
f:	275	72	30	7	5	2	1

χ^2-test as a test of goodness of fit (Conti...)

Problem 5: Four coins are tossed 160 times and the following results were obtained.

Numbers of heads:	0	1	2	3	4
Observed frequencies:	17	52	54	31	6

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

Test for independence attributes

The chi-square test can also be applied to test the association between attributes such as honesty, smoking etc, when the sample data is presented in the form of a contingency table with any number of rows and columns.

Contingency Table: A classification table containing *m* rows and *n* columns with observed frequencies is called a contingency table.

Test for independence attributes (Conti...)

Problem 1:The following table gives the classification of 150 workers according to sex and nature of work. Test whether the nature of work is independent of the sex of the work.

	Stable	Unstable
Males	60	30
Females	15	45

Solution: H_0 : The nature of the work is independent of the sex of the worker.

H_1 : The nature of the work is not independent of the sex of the worker.

Test for independence attributes (Conti...)

Degrees of freedom $=(m-1)(n-1)=(2-1)(2-1)=1.$

Level of significance = 5%

Table value of χ^2=3.84

Contingency Table:

	Stable	Unstable	Total
Males Females	60 15	30 45	90 60
Total	75	75	150

Expected frequencies are given in the following table

	Stable	Unstable
Males	75×90/150=45	75×90/150=45
Females	75×60/150=30	75×60/150=30

Expected Cell Frequency = (Row Total * Column Total)/N.

Calculation of χ^2

O_i	E_i	$(O_i - E_i)$	$\frac{(O_i - E_i)^2}{E_i}$
60	45	15	5.00
15	30	-15	7.50
30	45	-15	5.00
45	30	15	7.50
Total			25

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

The calculated value of χ^2 is greater than the table value at 5% LS and 1 df.
Hence, we reject the null hypothesis.

- Problem 2: A tobacco company claims that there is no relationship between smoking and lung ailments. To investigate the claim, a random sample of 300 persons in the age group of 40 and 50 are given a medical test. The observed sample results are tabulated below

	Lung ailment	Non-lung ailment
Smokers	75	105
Non smokers	25	95

On the basis of this information, can it be concluded that smoking and lung ailments are independent? ($\chi^2_{0.05} = 3.841$ for 1 df)