Chapter 6

* Goodness-of-Fit Test to determine if a population has a specified theoretical distribution.
* Independence Test to determine if there is association between the two variables in a contingency table where the data is all drawn from one population.
* Homogeneity Test to determine if two or more populations have the same distribution of a single categorical variable.

10.80 The grades in a statistics course for a particular semester were as follows:

Grade	A	B	C	D	F
f	14	18	32	20	16

Test the hypothesis, at the 0.05 level of significance, that the distribution of grades is uniform.

1. H_0: *The data follow uniform distribution.*

H_1: *The data doen't follow uniform distribution.*

We will use χ^2 test for goodness of fit.

Calculate Expected frequencies.

N=14+18+32+20+16=100

$e_i = Np_i \;,\quad p_i = \frac{1}{5} = 0.2 \;, i = 1, \dots, 5$

Thus, $E_i = 100(0.2) = 20 \;; \forall\, i$

Grade	A	B	C	D	F
Observed	14	18	32	20	16
Expected	20	20	20	20	20

2. Test statistic:

$$\chi^2 = \sum_{i=1}^{5} \frac{(O_i - e_i)^2}{e_i}$$

$$\chi^2 = \frac{(14-20)^2}{20} + \cdots + \frac{(16-20)^2}{20} = 10$$

3. Decision:

We Reject H_0 if $\chi^2 > \chi^2_{\alpha,k-1} = \chi^2_{0.05,4} = 9.488$
Since, $\chi^2 = 10 > 9.488$ we reject H_0 . i.e the data doesn't follow uniform distribution.

10.81 A die is tossed 180 times with the following results:

x	1	2	3	4	5	6
f	28	36	36	30	27	23

is this a balanced die? Use a 0.01 level of significance.

1. H_0: *The die is* balanced .

H_1: *The die is not* balanced .

We will use χ^2 test for goodness of fit.

Calculate Expected frequencies.

N=28+36+36+30+27+23=180 ;

$$e_i = Np_i \ , if\ the\ die\ is\ \ balanced\ \ p_i = \frac{1}{6}\ \ , i = 1, \ldots, 6$$

$$e_i = 180\left(\frac{1}{6}\right) = 30$$

x	1	2	3	4	5	6
Observed	28	36	36	30	27	23
Expected	30	30	30	30	30	30

2. Test statistic:

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

$$\chi^2 = \frac{(28-30)^2}{30} + \cdots + \frac{(23-30)^2}{30} = 4.47$$

3.Decision:

We Reject H_0 if $\chi^2 > \ \chi^2_{\alpha,n-1} = \chi^2_{0.01,5} = 15.086$
Since, $\chi^2 = 4.47 \ngtr 15.086$ we can't reject H_0. i.e the data is balanced.

10.87 A random sample of 90 adults is classified according to gender and the number of hours of television watched during a week:

	Gender	
	Male	Female
Over 25 hours	15	29
Under 25 hours	27	19

Use a 0.01 level of significance and test the hypothesis that the time spent watching television is independent of whether the viewer is male or female.

1. H_0: *Time spent watching TV independent of grades.*

(The two random variable are independent)

H_1: *The two random variable are dependent .*

We will use χ^2 test Independence Test

Calculate Expected frequencies.

2. Test statistic:

$$\chi^2 = \sum_{i,j} \frac{(O_{ij}-e_{ij})^2}{e_{ij}}; \qquad e_{ij} = \frac{\sum_i O_{ij} \sum_j O_{ij}}{N} = \frac{n_r n_c}{N}$$

	Gender		$\sum_j O_{ij} = O_{i\cdot} = n_r$
	Male	Female	
Over 25 hours	15	29	**44**
Under 25 hours	27	19	**46**
$\sum_i O_{ij} = O_{j\cdot} = n_c$	**42**	**48**	**N= 90**

$$e_{11} = \frac{42(44)}{90} = 20.53; \qquad e_{21} = \frac{42(46)}{90} = 21.47$$

$$e_{12} = \frac{48(49)}{90} = 23.47\,; \qquad e_{22} = \frac{48(46)}{90} = 24.53$$

$$\chi^2 = \frac{(15-20.53)^2}{20.53} + \frac{(27-21.47)^2}{21.53} + \frac{(29-23.47)^2}{23.47} + \frac{(19-24.53)^2}{24.53} = 5.47$$

3. Decision:

We Reject H_0 if $\chi^2 > \chi^2_{\alpha,(c-1)(r-1)} = \chi^2_{0.01,1*1} = \chi^2_{0.01,1} = 6.635$

Since, $\chi^2 = 5.47 \ngtr 6.635$ we can't reject H_0. i.e The two variables are independent.

10.89 A criminologist conducted a survey to determine whether the incidence of certain types of crime varied from one part of a large city to another. The particular crimes of interest were assault, burglary, larceny, and homicide. The following table shows the numbers of crimes committed in four areas of the city during the past year.

	Type of Crime				
District	**Assault**	**Burglary**	**Larceny**	**Homicide**	n_r
1	162	118	451	18	749
2	310	196	996	25	1527
3	258	193	458	10	919
4	280	175	390	19	864
n_c	1010	682	2295	72	4059

Can we conclude from these data at the 0.01 level of significance that the occurrence of these types of crime is dependent on the city district?

1. H_0: type of crime and district are independent .

H_1: type of crime and district are dependent .

We will use χ^2 test Independence Test

2. Test statistic:

$$\chi^2 = \sum_{i,j} \frac{(o_{ij}-e_{ij})^2}{e_{ij}}; \qquad e_{ij} = \frac{\sum_i o_{ij} \sum_j o_{ij}}{N} = \frac{n_r n_c}{N}$$

$$e_{11} = 186.37 \quad e_{12} = 125.85 \; e_{13} = 423.49 \quad e_{14} = 13.29$$

$$e_{21} = 379.96 \; e_{22} = 256.57 \quad e_{23} = 863.38 \quad e_{24} = 27.09$$

$$e_{31} = 228.67 \; e_{32} = 1544.41 \quad e_{33} = 519.61 \quad e_{34} = 16.3$$

$$e_{41} = 214.99 \; e_{42} = 145.17 \quad e_{43} = 988.5 \quad e_{44} = 15.38$$

$$\chi^2 = \frac{(186.37-162)^2}{186.37} + \cdots + \frac{(15.33-19)^2}{15.33} = 124.53$$

3.Decision:

We Reject H_0 if $\chi^2 > \chi^2_{\alpha,(c-1)(r-1)} = \chi^2_{0.01,3*3} = \chi^2_{0.01,9} = 21.666$
Since, $\chi^2 = 124.53 > 21.666$ we reject H_0. i.e The occurrence of these types of crim is dependent on the city district.

10.93 To determine current attitudes about prayer in public schools, a survey was conducted in four Virginia counties. The following table gives the attitudes of 200 parents from Craig County, 150 parents from Giles County, 100 parents from Franklin County, and 100 parents from Montgomery County:

	County				
Attitude	**Craig**	**Giles**	**Franklin**	**Mont.**	n_r
Favor	65	66	40	34	205
Oppose	42	30	33	42	147
No opinion	93	54	27	24	198
n_c	200	150	100	100	550

Test for homogeneity of attitudes among the four counties concerning prayer in the public schools. ~~Use a *P*-value in your conclusion.~~

1. H_0: $for\ each\ row\ i\ p_{i1} = \cdots = p_{ic}$.

H_1: $at\ least\ one\ of\ H_0\ statments\ is\ false$.

$\mathrm{H_0}$: For each attitude, the proportions of craig, Giles, Franklin and Montgomery counties are the same.

H_1: $At\ least\ one\ of\ the\ counties\ proportions\ is\ different.$

We will use χ^2 test Homogeneity Test.

2. Test statistic:

$$\chi^2 = \sum_{i,j} \frac{(o_{ij}-e_{ij})^2}{e_{ij}}; \qquad e_{ij} = \frac{\sum_i o_{ij} \sum_j o_{ij}}{N} = \frac{n_r n_c}{N}$$

$$e_{11} = 74.55 \quad e_{12} = 55.91 \quad e_{13} = 37.27 \quad e_{14} = 37.27$$

$$e_{21} = 53.45 \quad e_{22} = 40.09 \quad e_{23} = 26.73 \quad e_{24} = 26.73$$

$$e_{31} = 72.00 \quad e_{32} = 54.00 \quad e_{33} = 36.00 \quad e_{34} = 36.00$$

$$\chi^2 = \frac{(65-74.55)^2}{74.55} + \cdots + \frac{(24-36)^2}{36} = 31.1$$

3. Decision:

We Reject H_0 if $\chi^2 > \chi^2_{\alpha,(c-1)(r-1)} = \chi^2_{0.05,\ 2*3} = \chi^2_{0.05,\ 6} = 12.592$
Since, $\chi^2 = 31.1 > 12.592$ we reject H_0. i.e attitudes are not homogeneous.

10.94 A survey was conducted in Indiana, Kentucky, band Ohio to determine the attitude of voters concerning school busing. A poll of 200 voters from each of these states yielded the following results:

	Voter Attitude			
State	Support	Do Not Support	Undecided	n_r
Indiana	82	97	21	200
Kentucky	107	66	27	200
Ohio	93	74	33	200
n_c	282	237	81	**600**

At the **0.05** level of significance, test the null hypothesis that the proportions of voters within each attitude category are the same for each of the three states.

1. H_0: The proportions for each attitude of Indiana , kentucky, ohio are the same.

H_1: *At least one of the states proportions different.*

We will use χ^2 test Homogeneity Test

2. Test statistic:

$$\chi^2 = \sum_{i,j} \frac{(o_{ij} - e_{ij})^2}{e_{ij}} \quad ; \quad e_{ij} = \frac{\sum_i o_{ij} \sum_j o_{ij}}{N} = \frac{n_r n_c}{N}$$

$$e_{11} = 94 \qquad e_{12} = 79 \qquad e_{13} = 27$$

$$e_{21} = 94 \qquad e_{22} = 79 \qquad e_{23} = 27$$

$$e_{31} = 94 \qquad e_{32} = 79 \qquad e_{33} = 27$$

$$\chi^2 = \frac{(82-94)^2}{94} + \cdots + \frac{(33-27)^2}{27} = 12.56$$

3.Decision:

We Reject H_0 if $\chi^2 > \chi^2_{\alpha(,c-1)(r-1)} = \chi^2_{0.05,\ 4} = 9.488$

Since, $\chi^2 = 12.56 > 9.488$ we reject H_0. i.e attitudes are not homogeneous.