

Example 1 (Testing Goodness of Fit of a Discrete Distribution):

The results of 180 rolls of a die are given below:

X	Ox
1	28
2	36
3	36
4	30
5	27
6	23
TOTAL	180

Test if the die is fair.

Null hypothesis to be tested is

$H_0: P(X=x) = 1/6$ for $x = 1, 2, \dots, 6$

Read the data file DIE.csv in R.

```
> d1<-read.csv("g:/Stats24x7/R/DIE.csv",header=TRUE)
> attach(d1)
> names(d1)
[1] "X" "Ox"
> chisq.test(Ox)
```

chi-square test in R assumes EQUAL Expected Frequencies hence the probability vector under null does not need to be specified.

Chi-squared test for given probabilities

OUTPUT from R.

data: Ox
X-squared = 4.4667, df = 5, p-value = 0.4844

Example 2 (Testing Independence):

Given the following data, test if Voting Preference is independent of Gender (see worksheet Independence of the file [chi-square tests in excel.xlsx](#)).

GENDER	Voting Preferences			Row total
	Republican	Democrat	Independent	
Male	300	150	50	500
Female	250	300	50	600
Column total	550	450	100	1100

There are several ways to get the data table in R. One way is shown below"

```
> row1 <- c(300, 150, 50)
> row2 <- c(250, 300, 50)
> tbl <- rbind(row1, row2)
```

```
> tbl
  [,1] [,2] [,3]
row1 300 150  50
row2 250 300  50
```

```
> chisq.test(tbl)
```

Pearson's Chi-squared test

```
data: tbl
X-squared = 45.8333, df = 2, p-value = 1.115e-10
```

Example 3 (Testing Homogeneity of Proportions):

In a telephone survey, respondents were asked to indicate their level of agreement with the statement “Cigarette smoking should be banned in public places”. The results are shown in the table below: SA = strongly agree, A = agree, N = neutral, D = disagree, SD = strongly disagree. Test if there is no difference in Males and Females with respect to their level of agreement on the banning of smoking in public places.

Gender	SA	A	N	D	SD	TOTAL
F	40	38	16	37	5	136
M	16	25	11	25	11	88
TOTAL	56	63	27	62	16	224

(see worksheet Homogeneity of the file [chi-square tests in excel.xlsx](#)).

The null hypothesis H_0 : there is no difference in Males and Females with respect to their level of agreement on the banning of smoking in public places.

```
col1 <- c(40, 16)
col2 <- c(38, 25)
col3 <- c(16, 11)
col4 <- c(37, 25)
col5 <- c(5, 11)

tbl2 <- cbind(col1, col2, col3, col4, col5)

> tbl2
  col1 col2 col3 col4 col5
[1,]  40  38  16  37   5
[2,]  16  25  11  25  11

> chisq.test(tbl2)

Pearson's Chi-squared test

data:  tbl2
X-squared = 8.5748, df = 4, p-value = 0.07265
```