

# $\chi^2$ Test for Independence

January 17, 2021

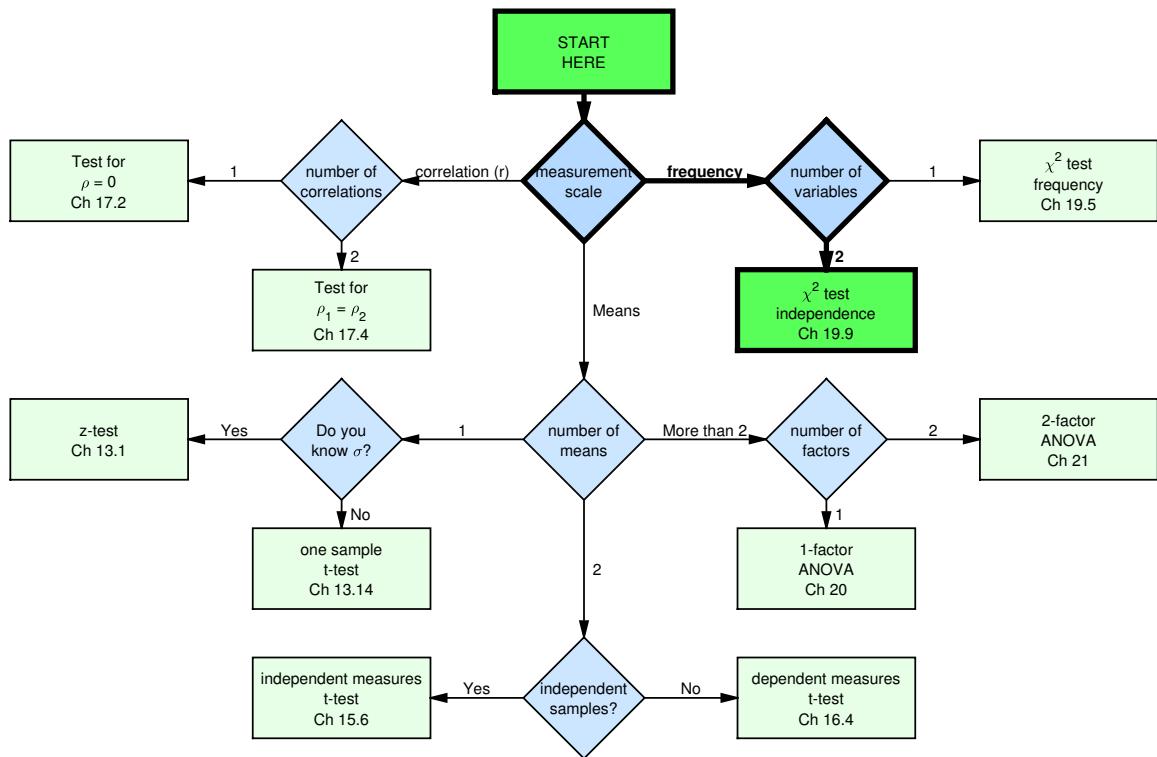
## Contents

- $\chi^2$  tests for independence
- Example 1: Computer users by gender
- Example 2: Does where you sit in class depend on gender?
- Using R to run a hypothesis test for independence
- Questions
- Answers

## $\chi^2$ tests for independence

This is the second kind of hypothesis test on frequency data. The test is to see if the frequency for which things fall into two variables are independent of each other. For example, you can use a  $\chi^2$  test to see if the choice of computers that students use varies with gender.

Here's how to get to the  $\chi^2$  test for independence with the flow chart:



This will make more sense with an example:

### Example 1: Computer users by gender

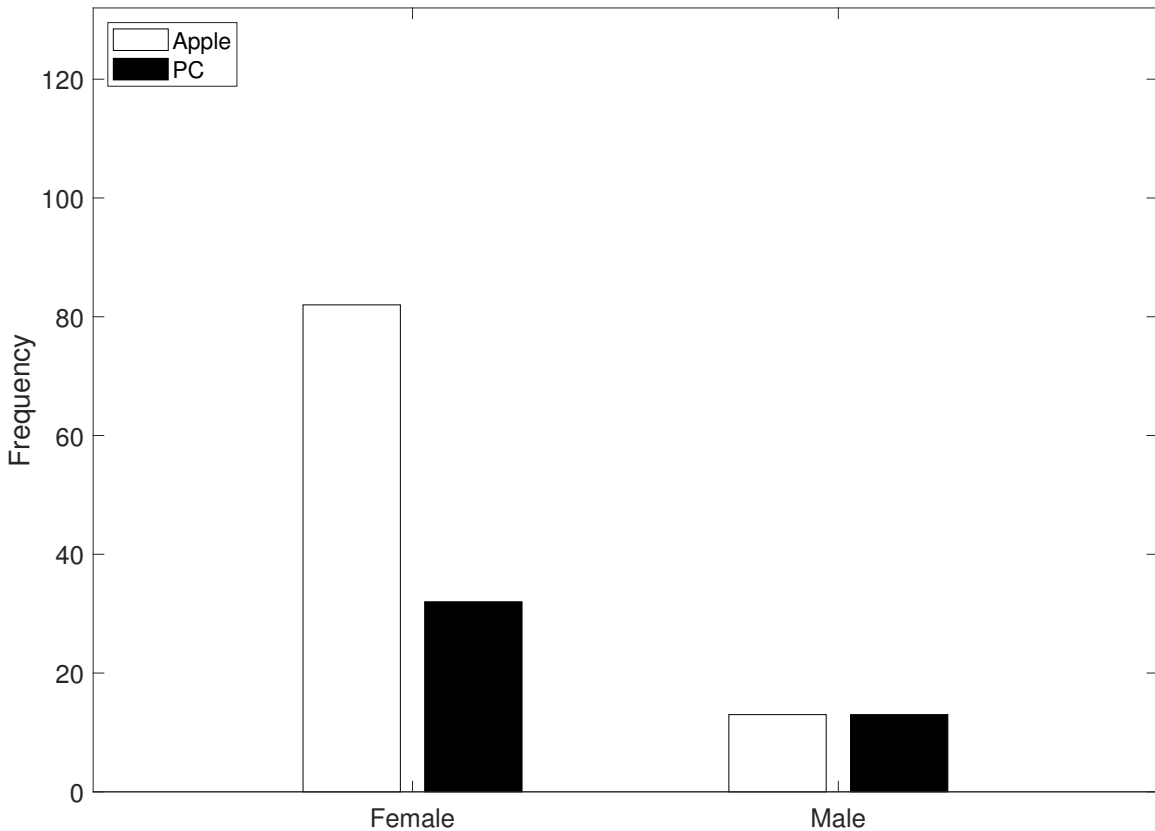
Does the type of computer that students use in our class depend on gender? To find out we'll run a  $\chi^2$  test for independence using  $\alpha = 0.05$ .

We'll use our survey data and count the number of students who use each kind of computer, depending upon their gender. This generates the following 2 X 2 table:

observed frequencies		
	Apple	PC
Female	82	32
Male	13	13

If you add up the rows and columns, you'll see that there are  $82 + 13 = 95$  students who use Apple computers, and  $32 + 13 = 45$  students who use PC computers. If computer choice did not depend on gender, then we should see a similar 95/45 ratio of computer use within the subsets of Female and Male students.

This type of frequency data can be visualized with a bar graph with a legend representing the second variable:



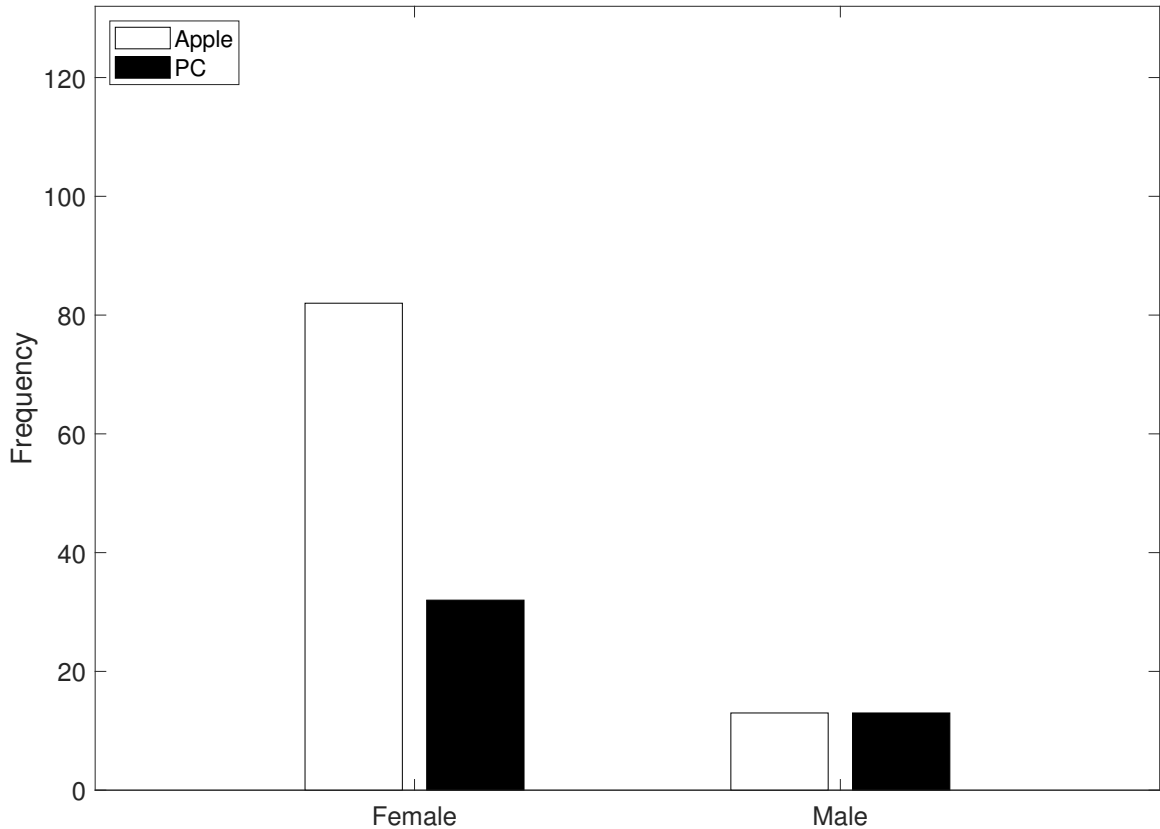
To test the null hypothesis that computer use does not depend on gender, we generate a table of expected frequencies. This table is set up so that the ratios of computer use is the same within each gender, but also ensures that the ratio of genders is the same across computer use.

An easy way to calculate the table of expected frequencies is to add up the rows and columns for observed frequencies. You then put the total number of observations in the bottom right cell of the table:

	Apple	PC	row sum
Female	82	32	114
Male	13	13	26
column sum	95	45	140

Then, for each cell, we multiply the sum for that cell's row with the sum for that cell's column and divide by the total number of observations:

Expected frequencies		
	Apple	PC
Female	$\frac{(114)(95)}{140} = 77.3571$	$\frac{(114)(45)}{140} = 36.6429$
Male	$\frac{(26)(95)}{140} = 17.6429$	$\frac{(26)(45)}{140} = 8.3571$



You should convince yourself that these expected frequencies have the right ratios under the null hypothesis. For example, for Female students, the expected ratio of Apple to PC computer users is  $77.3571/36.6429 = 2.11$ . For Male students, the ratio is the same:  $17.6429/8.3571 = 2.11$ . This is the same as the ratio of computer users for the whole class:  $95/45 = 2.11$

Going the other way, for Apple users, the expected ratio of Female to Male students is  $77.3571/17.6429 = 4.38$ . The ratio is the same for PC users:  $36.6429/8.3571 = 4.38$ . This is the same as the ratio of genders for the whole class:  $114/26 = 4.38$

Our hypothesis test depends on comparing these expected frequencies with the observed frequencies. Just like for the  $\chi^2$  test for frequencies, we calculate  $\chi^2$  with:

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

This time we sum across all cells, so that

$$\chi^2 = \frac{(82-77.3571)^2}{77.3571} + \frac{(13-17.6429)^2}{17.6429} + \frac{(32-36.6429)^2}{36.6429} + \frac{(13-8.3571)^2}{8.3571} =$$

$$0.2787 + 1.2218 + 0.5883 + 2.5794 = 4.6682$$

The degrees of freedom for this test is (number of rows - 1)(number of columns - 1) = df = (2-1)(2-1) = 1

You should see that this measure,  $\chi^2$ , is close to zero when the observed frequencies match the expected frequencies. Therefore, large values of  $\chi^2$  can be considered evidence against the null hypothesis of independence.

The critical value of  $\chi^2$  for df = 1 and alpha = 0.05 can be found in Table I in our book and also in the Excel spreadsheet:

df	0.995	0.99	0.975	0.95	0.9	0.1	0.05	0.025	0.01	0.005
1	0	0.0002	0.001	0	0.02	2.71	3.84	5.02	6.63	7.88
2	0.01	0.02	0.05	0.1	0.21	4.61	5.99	7.38	9.21	10.6
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

You can see that our observed  $\chi^2$  value (4.6682) falls above the critical value (3.84). We therefore reject the null hypothesis that the type of computer a student uses is independent on gender.

Just like for the t-table, the  $\chi^2$  table is not useful for calculating p-values. Instead we can use the  $\chi^2$ -calculator on the same page in the Excel spreadsheet, which gives us a p-value of 0.0307:

$\alpha$ to $\chi^2$		
df	$\alpha$	$\chi^2$
1	0.05	3.84

$\chi^2$ to $\alpha$		
df	$\chi^2$	$\alpha$
1	4.6682	0.0307

Using APA format, we'd state: "The choice of computer varies with gender,  $\chi^2(1, N = 140) = 4.6682, p = 0.0307$ ."

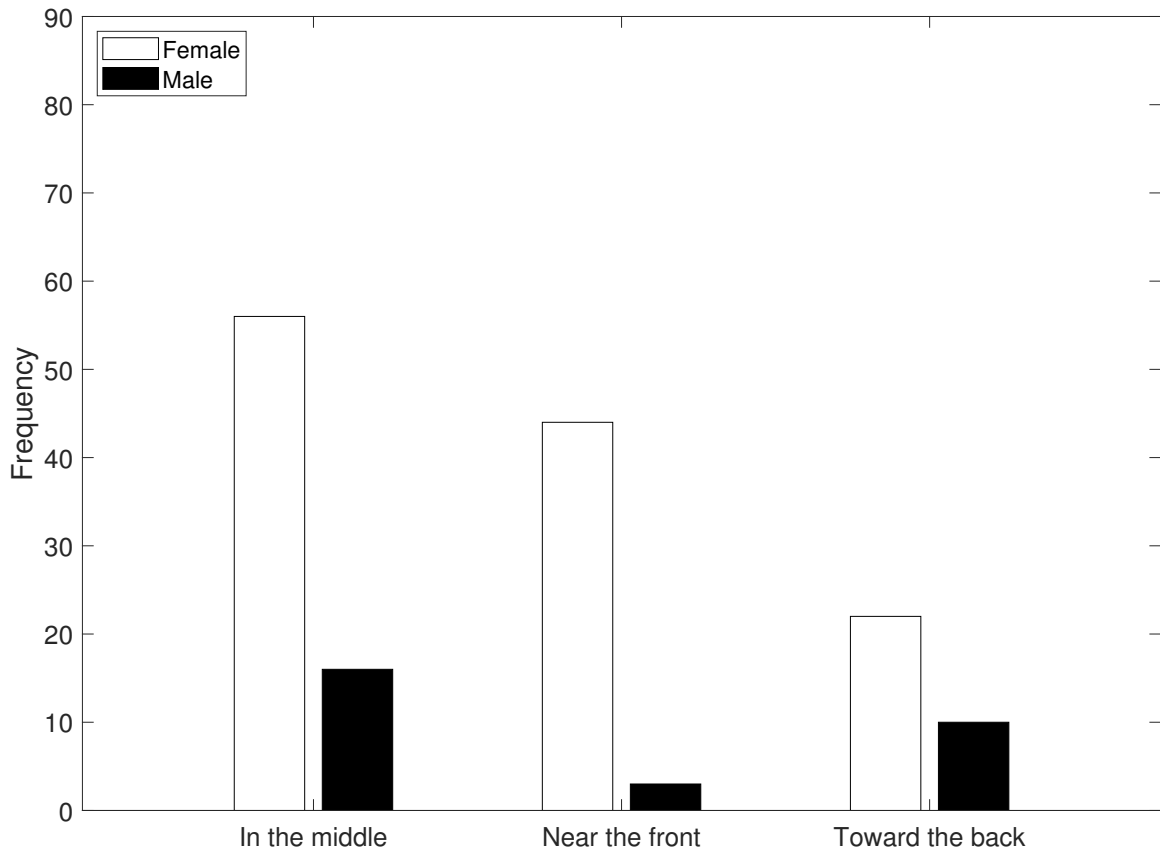
### Example 2: Does where you sit in class depend on gender?

Does your preference for where you sit in class depend on gender? We'll use our survey data to find out. This will be another  $\chi^2$  test for independence using  $\alpha = 0.05$ .

We first count the number of students across the 3 places to sit for each gender. This generates the following 3 X 2 table:

observed frequencies		
	Female	Male
In the middle	56	16
Near the front	44	3
Toward the back	22	10

Here's a bar graph of our frequency data:



Here are the row and column sums which are used to calculate the expected frequencies:

Sums and expected frequencies			
	Female	Male	sum
In the middle	$\frac{(72)(122)}{151} = 58.1722$	$\frac{(72)(29)}{151} = 13.8278$	72
Near the front	$\frac{(47)(122)}{151} = 37.9735$	$\frac{(47)(29)}{151} = 9.0265$	47
Toward the back	$\frac{(32)(122)}{151} = 25.8543$	$\frac{(32)(29)}{151} = 6.1457$	32
sum	122	29	151

Like the last example, we compare our observed and expected frequencies by calculating the  $\chi^2$  statistic:

$$\chi^2 = \frac{(56-58.1722)^2}{58.1722} + \frac{(44-37.9735)^2}{37.9735} + \frac{(22-25.8543)^2}{25.8543} + \frac{(16-13.8278)^2}{13.8278} + \frac{(3-9.0265)^2}{9.0265} + \frac{(10-6.1457)^2}{6.1457} =$$

$$0.0811 + 0.9564 + 0.5746 + 0.3412 + 4.0236 + 2.4172 = 8.3941$$

The degrees of freedom for this test is (number of rows - 1)(number of columns - 1) = df = (3-1)(2-1) = 2

The critical value of  $\chi^2$  for df = 2 and alpha = 0.05 can be found in Table I in our book and also in the Excel spreadsheet:

df	0.995	0.99	0.975	0.95	0.9	0.1	0.05	0.025	0.01	0.005
1	0	0.0002	0.001	0	0.02	2.71	3.84	5.02	6.63	7.88
<b>2</b>	<b>0.01</b>	<b>0.02</b>	<b>0.05</b>	<b>0.1</b>	<b>0.21</b>	<b>4.61</b>	<b>5.99</b>	<b>7.38</b>	<b>9.21</b>	<b>10.6</b>
3	0.07	0.11	0.22	0.35	0.58	6.25	7.81	9.35	11.34	12.84
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

You can see that our observed  $\chi^2$  value (8.3941) falls above the critical value (5.99). We therefore reject the null hypothesis that where you prefer to sit in class varies does not vary with gender.

The  $\chi^2$ -calculator on the same page in the Excel spreadsheet gives us a p-value of 0.015:

$\alpha$ to $\chi^2$		
df	$\alpha$	$\chi^2$
2	0.05	5.99



$\chi^2$ to $\alpha$		
df	$\chi^2$	$\alpha$
2	8.3941	0.015

Using APA format, we'd state: "The preference of where you prefer to sit varies across gender,  $\chi^2(2, N = 151) = 8.3941, p = 0.015$ ."

## Using R to run a hypothesis test for independence

The following R script covers how to run a  $Chi^2$  test for independence for the examples in this tutorial.

The R commands shown below can be found here: [Chi2TestIndependence.R](#)

```
# Chi-squared test for independence
#
# Just like for frequencies, R's 'chisq.test' provides a p-value for the chi-squared
# test for independence.

# Here we'll run the two examples in the chi_2_test_independence_tutorial

# Load in the survey data
survey <- read.csv("http://www.courses.washington.edu/psy315/datasets/Psych315W21survey.csv")
# Example 1: Computer users by gender

# create a table from the survey
fo <- table(survey$gender, survey$computer)
# Only keep the 1st (Apple) and 4th (PC) columns of the table
fo <- fo[c(2,3),c(1,3)]
# run the chi-squared test. The 'correct' parameter is a specific case for a
# 2x2 table. If 'correct = TRUE' then chisq.test will use a 'Yates continuity
# correction', which is essentially the same thing as we did back in the normal
# approximation to the binomial when we added or subtracted .5 from X. To get
# the same p-value as in the tutorial we'll set correct = FALSE:
out <- chisq.test(fo, correct = FALSE)
# The chi-squared statistic is:
out$statistic
X-squared
4.668103

# The degrees of freedom is:
out$parameter
df
1

# And the p-value is:
out$p.value
```

```
[1] 0.03072785
```

```
# Writing in APA format can be done like this:
```

```
sprintf('Chi-Squared(%d,N=%d) = %5.2f, p = %5.4f',out$parameter,sum(fo),out$statistic,out$p.value)
```

```
[1] "Chi-Squared(1,N=140) = 4.67, p = 0.0307"
```

```
# Plot the results:
```

```
barplot(fo,  
        beside=TRUE,  
        legend = row.names(fo))
```

```
# Example 2: Does where you sit in class depend on gender?
```

```
fo <- table(survey$gender,survey$sit)
```

```
fo <- fo[c(2,3),]
```

```
# Run the chi-squared test. If we don't specify the expected frequency the
```

```
# test assumes that expected frequencies are equal across categories.
```

```
out <- chisq.test(fo)
```

```
# result in APA format:
```

```
sprintf('Chi-Squared(%d,N=%d) = %5.2f, p = %5.4f',out$parameter,sum(fo),out$statistic,out$p.value)
```

```
[1] "Chi-Squared(2,N=151) = 8.39, p = 0.0150"
```

```
# Plot the results:
```

```
barplot(fo,  
        beside=TRUE,  
        legend = row.names(fo),  
        col = c("Pink","Blue"))
```

## Questions

Your turn. Here are 6 practice  $\chi^2$  tests for independence followed by their answers and R code.

1) Suppose statistics problems and professors come in 3 varieties: uneven, gamy and outgoing. Let's find 76 statistics problems and 63 professors and count how many fall into each variety. This generates the following table:

observed frequencies		
	statistics problems	professors
uneven	35	13
gamy	19	29
outgoing	22	21

Make a table of the expected frequencies.

Using an alpha value of  $\alpha=0.05$ , test the hypothesis that the statistics problems and professors are distributed independently across the varieties of uneven, gamy and outgoing.

2) Suppose neurons and personality disorders come in 3 varieties: bustling, decorous and tasteless. In your spare time you find 37 neurons and 24 personality disorders and count how many fall into each variety. This generates the following table:

observed frequencies		
	neurons	personality disorders
bustling	8	6
decorous	22	10
tasteless	7	8

Make a bar graph showing the frequencies for all varieties.

Make a table of the expected frequencies.

Using an alpha value of  $\alpha=0.05$ , test the hypothesis that the neurons and personality disorders are distributed independently across the varieties of bustling, decorous and tasteless.

3) Suppose planets and video games come in 2 varieties: parched and tender. In your spare time you find 57 planets and 50 video games and count how many fall into each variety. This generates the following table:

observed frequencies		
	planets	video games
parched	36	23
tender	21	27

Make a table of the expected frequencies.

Using an alpha value of  $\alpha=0.01$ , test the hypothesis that the planets and video games are distributed independently across the varieties of parched and tender.

4) Suppose nerds and statistics problems come in 2 varieties: past and angry. I find 33 nerds and 38 statistics problems and count how many fall into each variety. This generates the following table:

observed frequencies		
	nerds	statistics problems
past	25	20
angry	8	18

Make a table of the expected frequencies.

Using an alpha value of  $\alpha=0.05$ , test the hypothesis that the nerds and statistics problems are distributed independently across the varieties of past and angry.

5) Suppose professors and poo come in 2 varieties: deafening and lewd. Tomorrow you find 22 professors and 14 poo and count how many fall into each variety. This generates the following table:

observed frequencies		
	professors	poo
deafening	6	6
lewd	16	8

Make a table of the expected frequencies.

Using an alpha value of  $\alpha=0.01$ , test the hypothesis that the professors and poo are distributed independently across the varieties of deafening and lewd.

6) Suppose British people and personality disorders come in 2 varieties: loose and cruel. For some reason you find 34 British people and 26 personality disorders and count how many fall into each variety. This generates the following table:

observed frequencies		
	British people	personality disorders
loose	26	12
cruel	8	14

Make a table of the expected frequencies.

Using an alpha value of  $\alpha=0.05$ , test the hypothesis that the British people and personality disorders are distributed independently across the varieties of loose and cruel.

## Answers

1) The distribution of statistics problems and professors across the categories of uneven, gamy and outgoing.

Sums and expected frequencies			
	statistics problems	professors	sum
uneven	$\frac{(48)(76)}{139} = 26.2446$	$\frac{(48)(63)}{139} = 21.7554$	48
gamy	$\frac{(48)(76)}{139} = 26.2446$	$\frac{(48)(63)}{139} = 21.7554$	48
outgoing	$\frac{(43)(76)}{139} = 23.5108$	$\frac{(43)(63)}{139} = 19.4892$	43
sum	76	63	139

$$\chi^2 = \frac{(35-26.2446)^2}{26.2446} + \frac{(19-26.2446)^2}{26.2446} + \frac{(22-23.5108)^2}{23.5108} + \frac{(13-21.7554)^2}{21.7554} + \frac{(29-21.7554)^2}{21.7554} + \frac{(21-19.4892)^2}{19.4892} =$$

$$2.9209 + 1.9998 + 0.0971 + 3.5236 + 2.4125 + 0.1171 = 11.071$$

$$df = (3-1)(2-1) = 2$$

$$\chi_{crit}^2 = 5.99$$

We reject  $H_0$ .

The statistics problems and professors are not distributed independently of the varieties of uneven, gamy and outgoing,  $\chi^2(2, N=139)=11.071$ ,  $p=0.0039$ .

# Using R:

```
fo <- matrix(c( 35, 13, 19, 29, 22, 21), ncol = 3)
colnames(fo) <- c( "uneven", "gamy", "outgoing")
row.names(fo) <- c( "statistics problems", "professors")
fo
      uneven gamy outgoing
statistics problems    35    19     22
professors             13    29     21
out <- chisq.test(fo)
sprintf('Chi-Squared(%d,N=%d) = %5.2f, p = %5.4f', out$parameter, sum(fo), out$statistic, out$p.value)
[1] "Chi-Squared(2,N=139) = 11.07, p = 0.0039"
```

2) The distribution of neurons and personality disorders across the categories of bustling, decorous and tasteless.

Sums and expected frequencies			
	neurons	personality disorders	sum
bustling	$\frac{(14)(37)}{61} = 8.4918$	$\frac{(14)(24)}{61} = 5.5082$	14
decorous	$\frac{(32)(37)}{61} = 19.4098$	$\frac{(32)(24)}{61} = 12.5902$	32
tasteless	$\frac{(15)(37)}{61} = 9.0984$	$\frac{(15)(24)}{61} = 5.9016$	15
sum	37	24	61

$$\chi^2 = \frac{(8-8.4918)^2}{8.4918} + \frac{(22-19.4098)^2}{19.4098} + \frac{(7-9.0984)^2}{9.0984} + \frac{(6-5.5082)^2}{5.5082} + \frac{(10-12.5902)^2}{12.5902} + \frac{(8-5.9016)^2}{5.9016} =$$

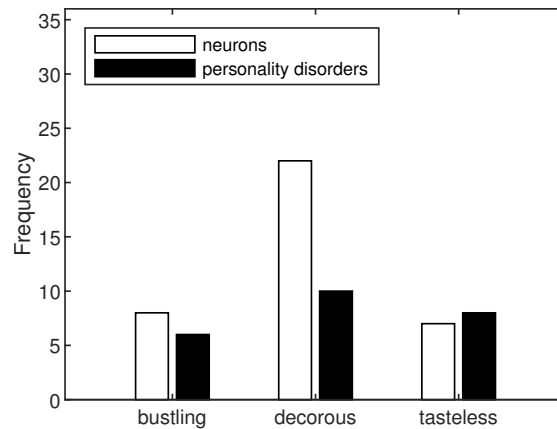
$$0.0285 + 0.3457 + 0.484 + 0.0439 + 0.5329 + 0.7461 = 2.1811$$

$$df = (3-1)(2-1) = 2$$

$$\chi_{crit}^2 = 5.99$$

We fail to reject  $H_0$ .

The neurons and personality disorders are distributed independently of the varieties of bustling, decorous and tasteless,  $\chi^2(2, N=61)=2.1811$ ,  $p=0.336$ .



# Using R:

```

fo <- matrix(c( 8, 6, 22, 10, 7, 8),ncol = 3)
colnames(fo) <- c( "bustling", "decorous", "tasteless")
row.names(fo) <- c( "neurons", "personality disorders")
fo

```

	bustling	decorous	tasteless
neurons	8	22	7
personality disorders	6	10	8

```

out <- chisq.test(fo)
sprintf('Chi-Squared(%d,N=%d) = %5.2f, p = %5.4f',out$parameter,sum(fo),out$statistic,out$p.value)
[1] "Chi-Squared(2,N=61) = 2.18, p = 0.3361"
# Plotting:
barplot(fo, beside = TRUE,
        legend = row.names(fo))

```



3) The distribution of planets and video games across the categories of parched and tender.

Sums and expected frequencies			
	planets	video games	sum
parched	$\frac{(59)(57)}{107} = 31.4299$	$\frac{(59)(50)}{107} = 27.5701$	59
tender	$\frac{(48)(57)}{107} = 25.5701$	$\frac{(48)(50)}{107} = 22.4299$	48
sum	57	50	107

$$\chi^2 = \frac{(36-31.4299)^2}{31.4299} + \frac{(21-25.5701)^2}{25.5701} + \frac{(23-27.5701)^2}{27.5701} + \frac{(27-22.4299)^2}{22.4299} =$$

$$0.6645 + 0.8168 + 0.7576 + 0.9312 = 3.1701$$

$$df = (2-1)(2-1) = 1$$

$$\chi_{crit}^2 = 6.63$$

We fail to reject  $H_0$ .

The planets and video games are distributed independently of the varieties of parched and tender,  $\chi^2(1, N=107)=3.1701, p=0.075$ .

# Using R:

```
fo <- matrix(c( 36, 23, 21, 27),ncol = 2)
colnames(fo) <- c( "parched", "tender")
row.names(fo) <- c( "planets", "video games")
fo
      parched tender
planets      36     21
video games  23     27
out <- chisq.test(fo)
sprintf('Chi-Squared(%d,N=%d) = %5.2f, p = %5.4f',out$parameter,sum(fo),out$statistic,out$p.value)
[1] "Chi-Squared(1,N=107) =  2.51, p = 0.1128"
```

4) The distribution of nerds and statistics problems across the categories of past and angry.

Sums and expected frequencies			
	nerds	statistics problems	sum
past	$\frac{(45)(33)}{71} = 20.9155$	$\frac{(45)(38)}{71} = 24.0845$	45
angry	$\frac{(26)(33)}{71} = 12.0845$	$\frac{(26)(38)}{71} = 13.9155$	26
sum	33	38	71

$$\chi^2 = \frac{(25-20.9155)^2}{20.9155} + \frac{(8-12.0845)^2}{12.0845} + \frac{(20-24.0845)^2}{24.0845} + \frac{(18-13.9155)^2}{13.9155} =$$

$$0.7976 + 1.3805 + 0.6927 + 1.1989 = 4.0697$$

$$df = (2-1)(2-1) = 1$$

$$\chi_{crit}^2 = 3.84$$

We reject  $H_0$ .

The nerds and statistics problems are not distributed independently of the varieties of past and angry,  $\chi^2(1, N=71)=4.0697$ ,  $p=0.0437$ .

# Using R:

```
fo <- matrix(c( 25, 20, 8, 18),ncol = 2)
colnames(fo) <- c( "past", "angry")
row.names(fo) <- c( "nerds", "statistics problems")
fo
      past angry
nerds      25    8
statistics problems  20  18
out <- chisq.test(fo)
sprintf('Chi-Squared(%d,N=%d) = %5.2f, p = %5.4f',out$parameter,sum(fo),out$statistic,out$p.value)
[1] "Chi-Squared(1,N=71) = 3.13, p = 0.0767"
```

5) The distribution of professors and poo across the categories of deafening and lewd.

Sums and expected frequencies			
	professors	poo	sum
deafening	$\frac{(12)(22)}{36} = 7.3333$	$\frac{(12)(14)}{36} = 4.6667$	12
lewd	$\frac{(24)(22)}{36} = 14.6667$	$\frac{(24)(14)}{36} = 9.3333$	24
sum	22	14	36

$$\chi^2 = \frac{(6-7.3333)^2}{7.3333} + \frac{(16-14.6667)^2}{14.6667} + \frac{(6-4.6667)^2}{4.6667} + \frac{(8-9.3333)^2}{9.3333} =$$

$$0.2424 + 0.1212 + 0.3809 + 0.1905 = 0.935$$

$$df = (2-1)(2-1) = 1$$

$$\chi_{crit}^2 = 6.63$$

We fail to reject  $H_0$ .

The professors and poo are distributed independently of the varieties of deafening and lewd,  $\chi^2(1, N=36)=0.935$ ,  $p=0.3336$ .

# Using R:

```
fo <- matrix(c( 6, 6, 16, 8),ncol = 2)
colnames(fo) <- c( "deafening", "lewd")
row.names(fo) <- c( "professors", "poo")
fo
```

```
      deafening lewd
professors      6  16
poo              6   8
```

```
out <- chisq.test(fo)
```

Warning message:

In `chisq.test(fo)` : Chi-squared approximation may be incorrect

```
sprintf('Chi-Squared(%d,N=%d) = %5.2f, p = %5.4f',out$parameter,sum(fo),out$statistic,out$p.value)
[1] "Chi-Squared(1,N=36) = 0.37, p = 0.5456"
```

6) The distribution of British people and personality disorders across the categories of loose and cruel.

Sums and expected frequencies			
	British people	personality disorders	sum
loose	$\frac{(38)(34)}{60} = 21.5333$	$\frac{(38)(26)}{60} = 16.4667$	38
cruel	$\frac{(22)(34)}{60} = 12.4667$	$\frac{(22)(26)}{60} = 9.5333$	22
sum	34	26	60

$$\chi^2 = \frac{(26-21.5333)^2}{21.5333} + \frac{(8-12.4667)^2}{12.4667} + \frac{(12-16.4667)^2}{16.4667} + \frac{(14-9.5333)^2}{9.5333} =$$

$$0.9265 + 1.6004 + 1.2116 + 2.0928 = 5.8313$$

$$df = (2-1)(2-1) = 1$$

$$\chi_{crit}^2 = 3.84$$

We reject  $H_0$ .

The British people and personality disorders are not distributed independently of the varieties of loose and cruel,  $\chi^2(1, N=60)=5.8313, p=0.0157$ .

# Using R:

```
fo <- matrix(c( 26, 12, 8, 14),ncol = 2)
colnames(fo) <- c( "loose", "cruel")
row.names(fo) <- c( "British people", "personality disorders")
fo
      loose cruel
British people      26      8
personality disorders  12     14
out <- chisq.test(fo)
sprintf('Chi-Squared(%d,N=%d) = %5.2f, p = %5.4f',out$parameter,sum(fo),out$statistic,out$p.value)
[1] "Chi-Squared(1,N=60) = 4.60, p = 0.0320"
```