Chi-squared tests in SPSS (Quiz)

The development of this E-Book has been supported by the British Academy.

Chi-squared test practical questions

Now you will explore whether there is an association between familiarity with a different environmental issue and degree of optimism about its prospects. The variable INFNUCL records how informed the participant felt about *nuclear waste* , and the variable OPTNUCL records whether they believed that the situation with nuclear waste would improve, stay the same, or get worse over the next 20 years.

Use the Crosstabs window to perform a chi-squared test on the variable **OPTNUCL** and **INFNUCL** and answer the following questions:

- Question: In the first output we see the numbers of observations and missing data. How many data points are used in the test?
- Question: Looking at the second output how many observations are there where **OPTNUCL** is Stay about the same and **INFNUCL** is Heard of, but not really able to explain this?
- Question: How many observations are expected under the null hypothesis of independence where **OPTNUCL** is Stay about the same and **INFNUCL** is Heard of, but not really able to explain this?
- Question: Looking at the third output does the asymptotic Chi-squared test suggest a significant relationship between the variables?

The development of this E-Book has been supported by the British Academy.

Solutions to Chi-squared test practical questions

The SPSS instructions are as follows:

- Select **Descriptive Statistics** from the **Analyze** menu.
- Select **Crosstabs...** from the **Descriptive Statistics** sub-menu.
- Click on the **Reset** button.
- Copy the **OPTNUCL** variable into the **Row(s):** box.
- Copy the **INFNUCL** variable into the **Column(s):** box.
- Click on the **Statistics...** button.
- In the window that appears click on the **Chi-square** tickbox.
- Click on the **Continue** button.
- Click on the **Cells...** button.
- In the window that appears:
- Under counts click on the **Expected** tick box to include it.
- Click on the **Continue** button.
- Click on the **OK** button.

- Question: In the first output we see the numbers of observations and missing data. How many data points are used in the test?

Solution: The output from SPSS is as follows:

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Nuclear waste will improve in 20 years * How informed about nuclear waste	4717	90.8%	477	9.2%	5194	100.0%

Here we see that there are 5194 observations of which 477 are missing resulting in 4717 that can be used in the test.

- Question: Looking at the second output how many observations are there where **OPTNUCL** is Stay about the same and **INFNUCL** is Heard of, but not really able to explain this?

Solution: The output from SPSS is as follows:

Nuclear waste will improve in 20 years * How informed about nuclear waste Crosstabulation

			How informed about nuclear waste				
			Never heard of this	Heard of, but not really able to explain this	Know something, could explain in general	Familiar with this, could explain well	Total
Nuclear waste will improve in 20 years	Improve	Count	36	208	278	202	724
		Expected Count	26.9	237.0	290.6	169.6	724.0
	Stay about the same	Count	59	554	644	321	1578
		Expected Count	58.5	516.5	633.3	369.7	1578.0
	Get worse	Count	80	782	971	582	2415
		Expected Count	89.6	790.5	969.2	565.7	2415.0
Total		Count	175	1544	1893	1105	4717
		Expected Count	175.0	1544.0	1893.0	1105.0	4717.0

There are 554 observations where **OPTNUCL** is Stay about the same and **INFNUCL** is Heard of, but not really able to explain this. This is out of a total of 1578 observations where **OPTNUCL** is Stay about the same and 1544 observations where **INFNUCL** is Heard of, but not really able to explain this.

- Question: How many observations are expected under the null hypothesis of independence where **OPTNUCL** is Stay about the same and **INFNUCL** is Heard of, but not really able to explain this?

Solution: The output from SPSS is as follows:

Nuclear waste will improve in 20 years * How informed about nuclear waste Crosstabulation

			How informed about nuclear waste				
			Never heard of this	Heard of, but not really able to explain this	Know something, could explain in general	Familiar with this, could explain well	Total
Nuclear waste will improve in 20 years	Improve	Count	36	208	278	202	724
		Expected Count	26.9	237.0	290.6	169.6	724.0
	Stay about the same	Count	59	554	644	321	1578
		Expected Count	58.5	516.5	633.3	369.7	1578.0
	Get worse	Count	80	782	971	582	2415
		Expected Count	89.6	790.5	969.2	565.7	2415.0
Total		Count	175	1544	1893	1105	4717
		Expected Count	175.0	1544.0	1893.0	1105.0	4717.0

Under the model of independence we expect to see 516.5 observations where **OPTNUCL** is Stay about the same and **INFNUCL** is Heard of, but not really able to explain this. This means there are 37.5 more observations where **OPTNUCL** is Stay about the same and **INFNUCL** is Heard of, but not really able to explain this than expected.

- Question: Looking at the third output does the asymptotic Chi-squared test suggest a significant relationship between the variables?

Solution: The output from SPSS is as follows:

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	24.286[a]	6	.000
Likelihood Ratio	24.056	6	.001
Linear-by-Linear Association	.000	1	.984
N of Valid Cases	4717		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 26.86.

The relationship between the categorical variables, **OPTNUCL** and **INFNUCL** was examined to look for associations. A chi-squared test with 6 degrees of freedom was performed resulting in a test statistic of 24.286. This results in an asymptotic p value which is .000 (reported as $p < .001$). and therefore we have strong evidence to reject the null hypothesis that **OPTNUCL** and **INFNUCL** are independent and there is therefore some relationship between the variables.

As with the greenhouse gas issue, there is a significant association between familiarity with the nuclear waste issue and optimism about how serious a problem it will be. There were too many uninformed participants who believed things would get better, and too many highly informed participants who believed things would get worse, for the constructs to be independent of one another.