

### 1.3 Tests for Independence

Data: Class Examples Flu/Vaccine

How to conduct tests for independence with SPSS is illustrated using the example from class.

#### Example 1

Do the data provide sufficient evidence that the vaccine is effective and helps preventing people from getting the flu?

		No Vaccine (0)	One Shot (1)	Two shots (2)
The data	Flu (1)	24	9	13
	No Flu (0)	289	100	565

Input Data to SPSS: Flu (Response), Vaccine (Explanatory): Data is in fluvaccine.sav (posted on website)

Use the following SPSS commands to create a cross-tab table for flu and vaccine, and perform a test of independence. Standardized residuals are calculated to identify cells of possible influence.

1. Data>Weight Cases  
Weight by Frequency Variable: Count  
OK
2. Analyze>Descriptive Statistics>Crosstabs  
Row(s): Flu  
Column(s): Vaccine  
Statistics popup: check Chi-square  
Cells popup: Counts: check Observed  
check Expected  
check Residuals: Adjusted standardized  
OK

Adjusted Standardized Residuals in SPSS are the Standardized Cell Residuals in Class Notes (Agresti text)

Flu * Vaccine Crosstabulation						
			Vaccine			Total
			0	1	2	
Flu	0	Count	289	100	565	954
		Expected Count	298.6	104.0	551.4	954.0
		Adjusted Residual	-3.1	-1.9	4.2	
	1	Count	24	9	13	46
		Expected Count	14.4	5.0	26.6	46.0
		Adjusted Residual	3.1	1.9	-4.2	
Total		Count	313	109	578	1000
		Expected Count	313.0	109.0	578.0	1000.0

Chi-Square Tests			
	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	17.313 <sup>a</sup>	2	.000
Likelihood Ratio	17.252	2	.000
Linear-by-Linear Association	14.915	1	.000
N of Valid Cases	1000		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 5.01.

Are catching the flu and vaccination status associated?

1.  $H_0$ : Flu and Vaccine are independent vs  $H_a$ :  $H_0$  not true, or  
 $H_a$ : Flu and Vaccine not associated vs  $H_a$ :  $H_0$  not true  
 $\alpha = 0.05$
2. Assume: Random sample, all  $\hat{\mu}_{ij} \geq 5$
3. Test statistic from output Pearson:  $X^2 = 17.313, df = 2$   
Likelihood Ratio:  $G^2 = 17.252, df = 2$
4. For both tests: P-value  $< 0.001$
5. Reject  $H_0$
6. At significance level of 5% the data provide sufficient evidence that getting vaccinated is associated with the vaccination status.

Is the vaccine preventing or giving the flu?

Adjusted standardized residuals  $\geq 1.96$  in absolute magnitude identify cells contributing to the significance.

See cells 00, 02, 20, and 22 in the output.

Interpretation of standardized residual for cell 00: The residual is negative and below -1.96. Significantly less people stayed flu free (Flu=0) after no shot (Vaccine=0) than would have been expected if vaccine and flu would be independent. Indicating that not getting vaccinated put you at higher risk of getting the flu.