

Comparing Measures of Association Doing Simulations in SPSS for Table 6

A variety of measures of association exist for tables. Knowing which measure of association to choose and how to interpret it can be challenging. This exercise uses the results in Table 6 of the earlier exercise and conducts a simulation to investigate a specific set of measures of association to see what they best indicate.

Tables can be represented in SPSS without having the microdata from which they were initially derived. To do this, the structure of the table is created using variables to index the categories that make up the cells in the table and an additional variable containing the cell counts, which is subsequently assigned as a weight variable. For example, Table 6 consists of four categories in the column variable representing the Atlantic Provinces and three categories in the row variable, which indicates a family’s financial forecast for the next year. The product of 4 categories by 3 is 12 cells. Thus, we create an SPSS data file with 12 cases – where each case represents a cell – and include separate variables for province and family-financial-forecast. A third variable is added to include the actual cell counts from the original table. Using the Weight Cases feature in SPSS, the third variable with cell counts is assigned as the weight variable. Running a crosstab that specifies province as the column variable and family-financial-forecast as the row variable will reproduce Table 6.

1. Start an SPSS session.
2. Once you have a clean Data Editor window in front of you, click on the “Variable View” tab at the bottom of this window and enter the information in Figure 1.

Figure 1

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	province	Numeric	8	0	Atlantic Province	{10, Nfld}...	None	8	Right	Scale
2	finsit	Numeric	8	0	Financial Situation	{1, Worse}...	None	8	Right	Scale
3	wght	Numeric	8	0	From unweighted SFS data	None	None	8	Right	Scale
4	ewght	Numeric	8	0	Expected Value	None	None	8	Right	Scale
5	d1wght	Numeric	8	0	Simulation 1 : monotonic increasing	None	None	8	Right	Scale
6	d2wght	Numeric	8	0	Simulation 2: monotonic decreasing	None	None	8	Right	Scale
7	d3wght	Numeric	8	0	Simulation 3: step down	None	None	8	Right	Scale
8	d4wght	Numeric	8	0	Simulation 4: mixed	None	None	8	Right	Scale
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										
21										
22										
23										
24										

The values for Province are 10 'Nfld', 11 'PEI', 12 'NS', 12 'NB' and for Finsit are 1 'Worse', 2 'Same', 3 'Better'.

- Next, click on the "Data View" tab and enter the data as shown in Figure 2

Figure 2 : Data View

	province	finsit	wght	ewght	d1wght	d2wght	d3wght	d4wght		var	var	var	var
1	10	1	72	76	672	1	224	134					
2	11	1	26	34	1	1	1	92					
3	12	1	105	103	1	1	1	367					
4	13	1	94	84	1	747	1	150					
5	10	2	356	309	1	1	225	270					
6	11	2	138	140	302	302	1	106					
7	12	2	413	421	916	916	1	184					
8	13	2	307	344	1	1	1	449					
9	10	3	246	289	1	672	225	270					
10	11	3	140	130	1	1	302	106					
11	12	3	400	394	1	1	916	367					
12	13	3	348	321	747	1	747	150					
13													
14													
15													
16													
17													
18													
19													
20													

Notice the assignment of the twelve cases representing each of the twelve cells in the original Table 6.

- To reproduce Table 6, select **Data / Weight Case** from the Data Editor menu. In the Weight Cases dialogue box, click the radio button for "Weight cases by frequency variable" and then select "wght" from the variable list and click on the arrow to move it to the weight variable box. Then click **OK**.
- From the menu, select **Analysis / Descriptive Statistics / Crosstabs...** and move the variable "province" to the Columns list and the variable "finsit" to the Rows list.
- Next, click on **Cells** and select Observed Counts and Column Percentages followed by Continue.
- From **Statistics**, select the following list: Chi-square, Correlations, Lambda, Gamma, Kendall's tau b and Eta. Continue and OK. Confirm that this matches Table 6 by comparing the table in the Output window with the original table in your other handout.

Record your findings in the table on the following page.

Table of Outcomes

	Original	Expected	Simulation 1	Simulation 2	Simulation 3	Simulation 4
Pearson Chi-square & sig.						
Lambda and sig.						
Eta						
Kendall's tau-b and sig.						
Gamma and sig.						
Spearman Correlation & sig.						
Pearson Correlation & sig.						

8. Complete the rest of this table by going to the Data Editor window, select **Data / Weight Cases** and replace the current weight variable with the next one. Then simply rerun the crosstabs command without resetting any of its options. Record the results in the table above and continue working your way across the different weight variables representing different distributions.

The results of the Expected weight variable show no relationship across the categories of province. Simulations 1 through 4 display distinct patterns in the table. Simulation 1 shows a monotonic increasing distribution across the categories of province; simulation 2 shows the opposite pattern; simulation 3 displays an equal distribution for only Nfld and constants for the other three provinces (this is a step-down distribution); and simulation 4 shows some very interesting differences across the four provinces but without a distinct monotonic or step pattern.

Which of the measures of association do you feel is the most informative, if any?