SSRIC Teaching Resources Depository
Public Opinion on Social Issues -- 1975-2004
Elizabeth N. Nelson and Edward E. Nelson, California State University, Fresno

Appendix D:
Computation of Measures of Association

There are many measures of association used to measure the strength of relationship. Each has advantages and disadvantages. In this module, we have used two--Cramer's V and Goodman and Kruskal's Gamma. This appendix describes how to compute these measures. However, you will use a statistical package such as SPSS for Windows to do the actual computations for the exercises.

Cramer's V is one of several measures based on chi square. Chi square itself is not a measure of association, but a test of the hypothesis that two variables are unrelated. V is equal to the square root of the following value--chi square divided by the product of the number of cases in the table and the smaller of two values--the number of rows minus one and the number of columns minus one. The chi square for table 3.2 was equal to 12.52. V would be equal to 12.52 divided by the square root of the product of 900 and 1 or 12.52 divided by 30 or 0.42.

Cramer's V should be used when one or both of the variables consist of an unordered set of categories. It varies from 0 to 1. The closer it is to 0, the weaker the relationship and the closer to 1, the stronger the relationship. V can never be negative.

Gamma assumes that both of the variables consist of ordered categories. To understand the computation of Gamma, you must think of pairs of cases. Imagine four individuals. We'll just call them A, B, C, and D. For each person, we know their income and education which has been categorized as low, medium, and high. The table below displays these values.

Education of Respondent			
Income	High	Medium	Low
High	A		
Medium		C	

Low	B		D

These four individuals form six possible pairs. A can be paired with B, A with C, A with D, B with C, B with D, and C with D. Notice that the AB pair is the same as the BA pair since they involve the same two individuals.

A has more education and income than C. This is what we call a **concordant pair**. A is higher on both variables (i.e., income and education) than C. The AD pair is also concordant. A is higher on both education and income than D. And the CD pair is also concordant. C has more education and income than D.

However, C has more income than B, but C has less education than B. This is what we call a **discordant pair**. C is higher on one variable (i.e., income), but lower on the other variable (i.e., education).

When we look at the AB pair we see another possibility. A has more income than B, but A and B are tied on education. This is what we call a tied pair. The BD pair is also tied. B has more education than D, but B and D have the same income.

Gamma ignores all tied pairs. Since two of the six possible pairs are tied, Gamma would be based on the remaining four untied pairs. Gamma is equal to the number of concordant pairs (C) minus the number of discordant pairs (D) divided by the sum of the number of concordant pairs and discordant pairs.

In this example, Gamma would equal (3-1)/(3+1) or 2/4 or 0.50. Since there are more concordant pairs than discordant pairs, we can observe that it is more common for pairs to have the same order on both variables than to have different orders. In other words, large amounts of education tend to go with large amounts of income and small amounts of education tend to go with small amounts of income. This is what we call a **positive relationship**. Gamma has a positive sign if the relationship is positive [Note 1].

If the number of discordant pairs had been greater than the number of concordant pairs, the relationship would have been **negative** and the sign of Gamma would have been negative. This would have meant that large values of one variable would tend to go with small values of the other variable and that small values of one variable would tend to go with large values of the other variable.

The numerical value of Gamma tells us the strength of the relationship. The closer the value of Gamma to 0, the weaker the relationship and the closer to 1, the stronger the relationship [Note 2].

Notes

1. This will be true only if the columns are arranged from high to low (left to right) and the rows are arranged from high to low (top to bottom).

2. The fact that Gamma ignores all tied pairs tends to inflate the value of Gamma. For this reason, Gamma produces a larger measure of association than other measures.

REFERENCES AND SUGGESTED READING

- Herman J. Loether and Donald G. McTavish, 1993, *Descriptive and Inferential Statistics An Introduction* (4th edition), Allyn and Bacon.
- Marija J. Norusis, 2005, *SPSS 13.0 Guide to Data Analysis*, Prentice Hall.
