POLI 443
Applied Political Research

Session 8: Crosstab Analysis

Lecturer: Prof. A. Essuman-Johnson, Dept. of Political Science
Contact Information: aessuman-johnson@ug.edu.gh

UNIVERSITY OF GHANA
College of Education
School of Continuing and Distance Education
2014/2015 – 2016/2017

Bivariate Data Analysis: Cross Tabulations

- **Introduction**
- The procedure for measuring relationships and testing hypotheses depends on the level of measurement of the independent and dependent variables. When the independent and dependent variables are both nominal or ordinal level measures, contingency table analysis or cross tabulation is used. This unit will consist of the following sections.

Cross tabulation

- A cross tabulation or crosstab is a table that takes each case of a set of data and displays the value of each case for the two variables. This is done by putting the values for one variable along one side of the table and the value of for the other variable along the other side of the table. Each case is then placed in the cell in the table that corresponds to the case's values for both variables.

Illustration

- A researcher is interested in testing the hypothesis that Ghanaian farmers are more likely to vote NDC than University students in general elections. Data was collected on how farmers and students voted in a general election for a sample of voters. The first 9 cases in the sample was as follows:

Case No.	Occupation	Party Vote
1	Farmer	NDC
2	Farmer	NPP
3	Farmer	NDC
4	Student	NPP
5	Student	NDC
6	Farmer	NDC
7	Student	NPP
8	Student	NPP
9	Student	NPP

- A crosstab showing each case's value for both variables is done by putting the independent variable across the top of the table and the dependent variable down the left hand side (this is the conventional way to do it).

Table 1

Crosstab of the Relationship between Occupation and Party Vote

Dependent Variable: Party Vote	Independent Variable: Occupation		Total
	Farmer	Student	
a) NDC	1, 3, 6	5	
NPP	2	4, 7, 8, 9	
a) NDC	3	1	4
NPP	1	4	5
Total	4	5	9

How does table 1(b) help the researcher to measure the relationship and test the hypothesis that farmers vote differently from students? The expectation is that farmers will be more likely to vote NDC than students.

Of the 4 farmers in the table, 3 or 75% of them voted NDC. Of the 5 students 1 or 20% voted NDC. Therefore a greater proportion of farmers voted NDC than did students. This indicates that farmers and students vote differently. Thus, the value of the independent variable (farmer or student) matters and knowing a case's status would help us to account for the case's party vote.

Table 2

Crosstab of the Relationship between Occupation and Party Vote (Actual Numbers)

Dependent Variable: Party Vote	Independent Variable: Occupation		Total
	Farmer	Student	
a) NDC	300	400	700
NPP	200	600	800
Total	500	1,000	1,500
a) NDC	60%	40%	47%
NPP	40%	60%	53%
Total	100%	100%	100%

Looking at Table 2 we can assess that existence, direction and strength of the relationship between occupation and party vote.

First does a relationship exist between occupation and party vote? That is does the votes of farmers differ from votes of students? To do this we use percentages in each column. Of the 500 farmers who voted, 300 or 60% voted NDC, 200 or 40% voted NDC. Of the 1000 students who voted, 400 or40% voted NDC and 600 or 60% voted NPP. The percentage table in Table 2 shows at a glance that farmers vote much more for NDC than students.

- We can conclude from this that the two groups differ on the independent variable and thus there is a relationship between occupation and party vote. Table 3 is a crosstab showing no relationship between the independent and dependent variables. The cases in all the categories of the independent variable behaved the same on the dependent variables.

Table 3

- **Crosstab of the Relationship between Occupation and Party Vote showing No Relationship between Variables**

Dependent Variable: Party Vote	Independent Variable: Occupation		Total
	Farmer	Student	
NDC	47%	47%	47%
NPP	53%	53%	53%
Total	100%	100%	100%
N	(500)	(1,000)	1,500

- The percentage of cases with a particular value of the independent variable is the same for every category of the independent variable: an equal proportion of farmers and students voted NDC and NPP. As a result the hypothesis that occupation affects party vote would not be confirmed by this evidence. The categories of the independent variable may be either the rows or columns in a crosstab.

The convention is to place the independent variable across the top of a crosstab and thus creating the categories of the independent variable in the columns of the table. The variable with more categories can be placed down the left hand side of the table. It does not really matter whether a crosstab is constructed with the independent variable across the top or down the left hand side.

Summary

- In this session we have learned about how to construct cross-tabulations and use them to analyze the relationships between independent and dependent variables.