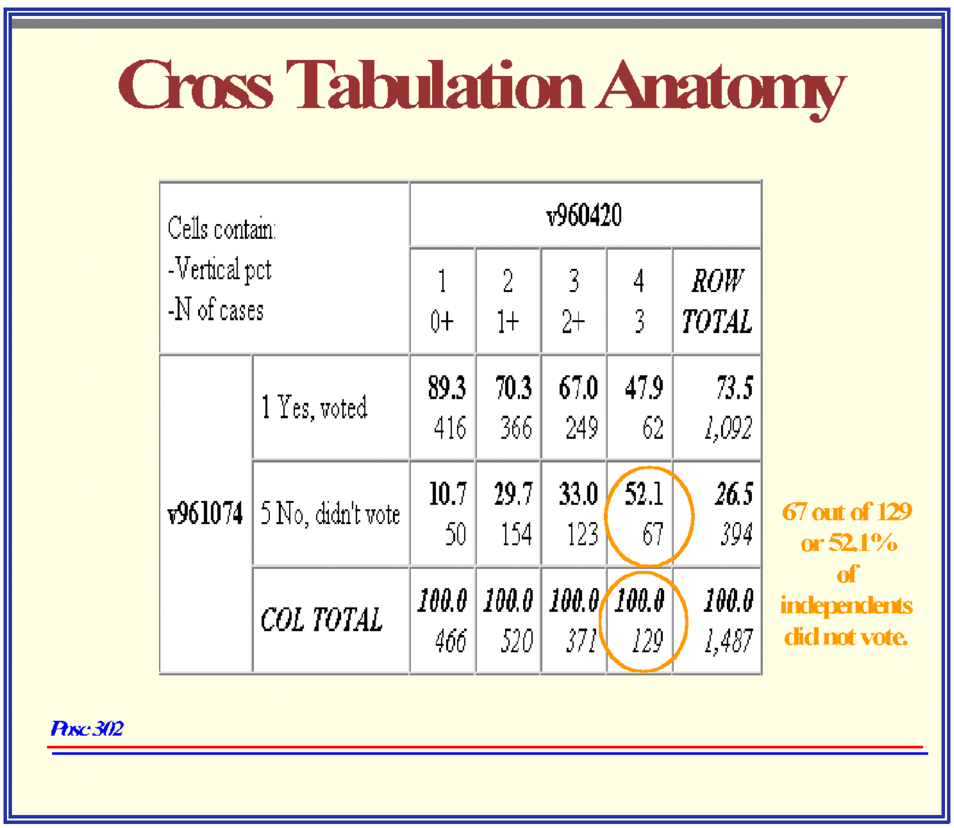
DEPARTMENT OF POLITICAL SCIENCE
AND
INTERNATIONAL RELATIONS
Research Methods Posc 302

RESEARCH DESIGN

I. TODAY'S SESSION:
- A. Surveys
  - 1. Cross-classifications
  - 2. Interpretations: what can be concluded
- B. Experiments
  - 1. Causal attributions

II. CROSS-CLASSIFICATIONS:
- A. Interpretation of cell entries.
  - 1. Example from last time:



Cells contain: -Vertical pct -N of cases		v960420				
		1 0+	2 1+	3 2+	4 3	*ROW TOTAL*
v961074	1 Yes, voted	89.3 416	70.3 366	67.0 249	47.9 62	*73.5* *1,092*
	5 No, didn't vote	10.7 50	29.7 154	33.0 123	52.1 67	*26.5* *394*
	COL TOTAL	*100.0* *466*	*100.0* *520*	*100.0* *371*	*100.0* *129*	*100.0* *1,487*

Figure 1: Voting by Partisanship

2. Table total **N**, the number of "cases" in the table.
   i. In the example above N = 1,487.
   ii. The number of cases in each category of the variables.
3. Table total **N**, the number of "cases" in the table.
   1) There are **row** and **column** marginal totals.
   2) They indicate the number of cases in each value of the variable.
   3) The column marginal totals divided by the total number of cases, N, give the percent in each category.
   4) So, for instance, 1,092 out of a total of 1,487 claimed to have voted; that is, 73.5 percent.
4. Cell frequencies: the number of cases having value i on the dependent variable ***and*** value j on the independent variable where "i" means the ith value of the dependent variable and j means the jth value of the independent variable.
   i. The first cell of the table is denoted by i = 1 and j = 1.
   ii. The second cell in the first row is denoted by i = 1 and j = 2.
   iii. Similarly, the 4^{th} cell in the second row is indicated by i = 2 and j = 4.
5. The entries in the body of the table are the **column** percentage and number of cases in each combination of categories.
   i. Example: 416 are code 1 on partisanship ***and*** 1 on voted.
      1) In the notation N_{11} = 416.
   ii. Similarly, 366 are coded 2 on partisanship ***and*** 1 on voted.
      1) In the notation N_{12} = 366.
   iii. The entries are just the number of occurrences of each combination of variable values.
6. **Percentages**: in this table cell entries are percentages, which the number in a category or combination of categories divided by a total and multiplied by 100.
   i. Percentage:

$$\boxed{Percent = \left(\frac{Cell\ frequency}{category\ total}\right)100 = \left(\frac{N_{ij}}{N_{+j}}\right)100}$$

Percentage calculation

   ii. The notation just follows the ideas presented above.
   iii. The column percent in the first cell of the table is just $(N_{11}/N)100$ =

416/466*100 = 89.3%.

iv. That is, we see that about 89.3 percent of the strong partisan respondents report that they voted.

v. Another example: look at the second column, second row: the number tells how many people in the second category of partisanship said they didn't vote: N_{22} = 154.

1) This is (154/520)100 = 29.7%

B. It is extremely important that you keep percentages straight.

1. Tip: if you see a percent, ask this is a percent of what?

i. Example: look at the percent in the second row and fourth column of the table.

ii. It's 52.1.

iii. What does that mean? It means that 52.1 percent of the 129 people in the last column of the horizontal (independent) variable did not vote.

III. INTERPRETATION:

A. Using percentages look for patterns.

1. Percentages can tell you how one variable is related to another.

2. More precisely, they can indicate how categories of one variable are related to categories of the other variable.

B. More examples:

1. Let first take some artificial cases. Here's a table that shows how party identification is related to attitudes toward President Clinton

Attitude/Party	Democrats	Republicans	Totals
Favor Clinton	100.0 60	0.0 0	60
Oppose Clinton	0.0 0	100.0 75	75
Totals	100 60	100. 75	135

i. It indicates that 60 out of 60 Democrats (or 100 percent) favor the president, where as 75 out of 75 (100%) Republicans oppose him.

1) You can figure out the meaning with just little thought.

ii. The data suggest that party is strongly related to attitudes, as we might expect.

2. Now consider another example.

Race/Turnout	White	African-American	Totals
Voted	70.0 140	70.0 49	189
Did not Vote	30.0 60	30.0 21	81
Totals	100.0 200	100.0 70	200

i. The percent of whites who report that they voted (70%) is the same as the percent of African-Americans who voted.

ii. Consequently, race doesn't distinguish voters and non-voters.

iii. There is, in other words, no relationship between turnout and race, at least in this hypothetical data set.

C. If one variable is associated with another one, the some categories of the first will tend to "go with" some categories of the other.

1. Otherwise there will be no clearly discernible pattern.

2. Detecting a pattern is partly a matter of judgment.

3. We'll see later that there are some statistical indices that can help us make the judgment.

D. The **strength of relationship**:

1. If categories of one variable seem to be associated with certain categories of another, we say the variables are strongly related.

2. Example 1:

i. Note that this table contains **only** percentages.

Party/Opinion	Democrats	Independents	Republicans
More money for education	90.0	10.0	5.0
Same amount for education	5.0	80.0	15.0
Less money for education	5.0	10.0	80.0
Totals	100.0	100.0	100.0

ii. Here the variable party identification is strongly related to opinion on federal spending for education.

iii. Why? Because if we know a person's party preference we can predict fairly accurately the person's opinion on the spending question.

1) 90 percent of Democrats favor more spending, whereas only 10 percent of independents and 5 percent of Republicans.

2) And conversely, only 5 percent of Democrats want to cut federal spending while 80 percent of the Republicans do.

3) Thus, for this hypothetical case party identification is strongly associated with opinions on educational spending.

3. Example 2:

i. Again the table contains **only** percentages.

ii. It relates partisanship to attitudes toward the death penalty.

Party/Opinion	Democrats	Independents	Republicans
Favor death penalty	60.0	55.0	65.0
Not sure about death penalty	20.0	25.0	10.0
Against death penalty	20.0	20.0	25.0
Totals	100.0	100.0	100.0

iii. In this case party is at most weakly related to the variable attitude toward capital punishment.

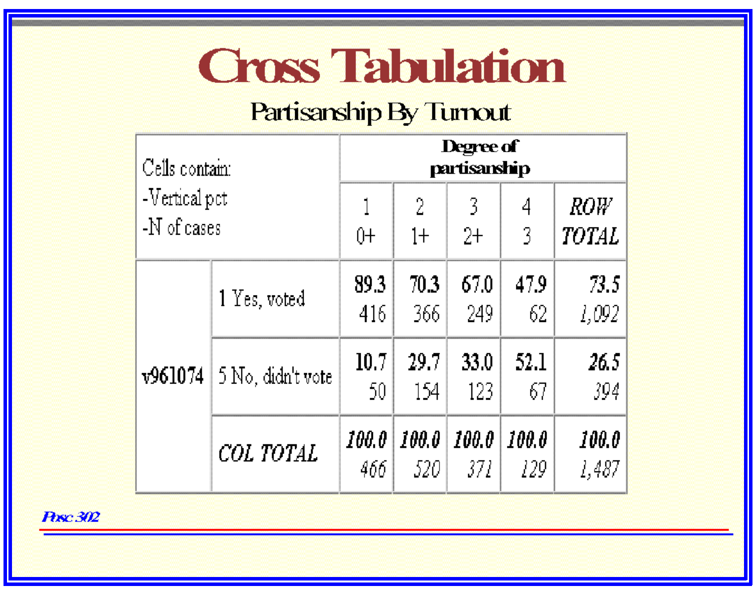
iv. Note, for instance, that about equal percentages of Democrats, independents, and Republicans are in favor.

v. Knowing a person's party doesn't help predict that person's on the issue.

4. Example 3:

i. Let return to the data previously analyzed, the table relating partisanship and turnout.

ii. Here is the table once again.



Cross Tabulation

Partisanship By Turnout

Cells contain: -Vertical pct -N of cases		Degree of partisanship 1 0+	2 1+	3 2+	4 3	ROW TOTAL
v961074	1 Yes, voted	89.3 416	70.3 366	67.0 249	47.9 62	73.5 1,092
	5 No, didn't vote	10.7 50	29.7 154	33.0 123	52.1 67	26.5 394
	COL TOTAL	100.0 466	100.0 520	100.0 371	100.0 129	100.0 1,487

Figure 2: Voting by Partisanship

iii. The percentages seem to support the argument that degree of partisanship is related to voting.
1) In particular the weaker the partisanship, the less likely a person will have voted.
2) We might express this idea with a simple drawing.

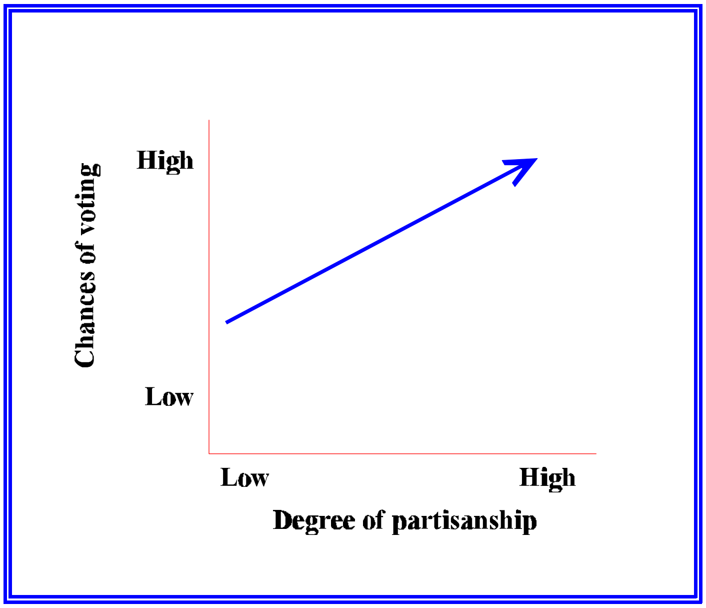


Figure 3: Interpretation of Relationship

3) Interpretation: as partisanship increases so does likelihood of turnout.
4) This is a verbal interpretation of the numbers in the table.

IV. CAUSATION:
A. So far we've talked about one variable being **related** to another.
1. Among others things surveys are useful tools for finding such relationships.
2. But they are less helpful if we want to make stronger statements.
B. Consider these claims:
1. "The decline in political parties in the United States has **led to** a decline in political participation, especially among members of the lower classes."
i. What is being asserted? One thing, the decline of parties, has **caused** another thing, the lowering of turnout.
2. "The death penalty **deters** crime."
i. The claim here is not simply that the presence or absence of the death penalty is associated with crime rates.
ii. Rather something stronger is claimed, namely that the existence of a

death penalty provision in an area (e.g., state) **causes** a reduction in crime.

3. "Generous welfare programs encourage idleness and irresponsible behavior."
    i. See, for example, Charles Murray, ***Losing Ground***.
    ii. The argument in this instance is that government welfare programs (i.e., public assistance like food stamps, medicaid) encourages or **causes** recipients to refuse to accept jobs, learn skills, or whatever else it takes to be responsible citizens.
4. "We need to maintain a strong defense because military weakness incites aggression."
    i. Politicians love to say that unless we have a large, modern military establishment potential aggressors will interpret our action as timidity and lack of resolve, miscalculate, and then engage in aggressive actions.
    ii. Put simply, lack of military preparedness **causes** aggression by our enemies.
5. Finally, "Pornography causes violence against women."

C. Causal attributions.
1. Each of these statements makes a causal claim: X causes Y.
    i. That is, the existence or occurrence of X leads to the existence or occurrence of Y.
    ii. Sometimes the idea is expressed a bit differently: If X takes place, Y will necessarily follow.
    iii. Or, changing the status of X inevitably leads to a change in the status of Y.
2. A common but not always good way of expressing the same idea is: X is a necessary and sufficient condition for Y.

D. The problem:
1. However stated a causal attribution is a strong claim.
    i. It asserts more than just X and Y are associated or appear together.
    ii. Rather it indicates a direct physical connection.
2. One might illustrate the difference between a causal attribution and a statement of association or relationship as in the next figure.

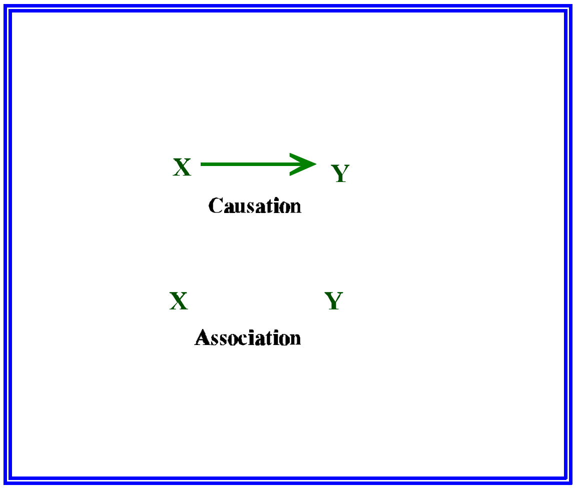


Figure 4: Causation Versus Association

E. Example:

1. Consider two variables:
   - i. Dependent: birth rate (number of live births per 1,000 women aged 18 and older) in various counties.
   - ii. Independent: number of storks per square mile.
2. It possible that we might observe an association such as the following.

Storks/births	Fewer than 10	11 to 49	50 or more
Fewer than 5	80.0%	20.0%	5.0%
6 to 9	15.0%	60.0%	10.0%
More than 10	5.0%	20.0%	85.0%
Totals	100.0%	100.0%	100.0%

.

   - i. This is just a simple contingency table or cross-classification that suggests that the greater the number of storks (per square mile) the greater or higher the birth rate in counties.
   - ii. But would these data support a causal attribution that storks are somehow causally involved in births and that the more of them around the more births are therefore possible?

F. A more substantive example:

1. Does partisanship have a causal impact on turnout?
   i. That is, does partisanship force people to participate at certain rate in the same way as a rise in temperature forces an increase in pressure in a contain of a given size?

G. The “classical” randomized, controlled scientific experiment throws light on these issues.

V. NEXT TIME:

A. Experimental designs

B. Approximations to experiments
   1. Quasi or naturalistic experiments.
   2. How survey data can be “manipulated” to permit causal inferences in non-experimental settings.

C. Reading:
   1. Johnson and Joslyn, ***Research Methods***, Chapter 5.