Crosstabs

- Independent Variables in columns
- Dependent Variables in rows
- To see what the chances are of committing a type I error, you perform a chi square test of independence
- Makes few assumptions about the data

Chi Square

- To check the probability that you’re committing a Type I Error, perform a Chi Square Test
- Once you get a Chi Square value, see if this value is statistically significant.

What a significant chi square value tells us

- We like significant chi square values. They tell us that the probability that our finding arose by chance is relatively slim.
- Significant is p<.1 (There is less than a 10% chance that we are committing a Type I error.)
What it does not tell us

- That the sample is free of bias,
- That the relationship is strong
- That the research design is strong
- That the measures are reliable and valid.

Measures of Association

- We still don’t know how strong a relationship is—that’s where measures of association come in.
- A statistic that indicates the strength of the relationship between a DV and an IV
- How strongly do two variables “go together.”

Choosing a measure of association

- To choose the proper measure of association, decide if you’ve got nominal, ordinal, or interval data

Choosing the measure of association

- Nominal data: Lambda and Cramer’s V
- Ordinal data: Kendall’s tau
- Interval data: Pearson’s $r$
Values of measures of association

- Nominal measures of association vary between 0.00 and 1.00.
- Ordinal and interval measures of association vary between -1.00 and +1.00.

Strong and Weak Relationships

- For Nominal Data
  - The closer the measure of association is to 0.00 the weaker the association
  - The closer the measure of association is to 1.00 the stronger the association
  - One is a “perfect” association

Percent Reduction in Error

- It indicates how much the knowledge of the distribution of the independent variable reduces the error in predicting the distribution of the dependent variable

Gender and the 2000 Election

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bush</td>
<td>900</td>
<td>200</td>
<td>1,100</td>
</tr>
<tr>
<td>Gore</td>
<td>100</td>
<td>800</td>
<td>900</td>
</tr>
<tr>
<td>Total</td>
<td>1,000</td>
<td>1,000</td>
<td>2,000</td>
</tr>
</tbody>
</table>
Gender and the 2000 Election

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bush</td>
<td>900</td>
<td>200</td>
<td>1,100</td>
</tr>
<tr>
<td>Gore</td>
<td>100</td>
<td>800</td>
<td>900</td>
</tr>
<tr>
<td>Total</td>
<td>1,000</td>
<td>1,000</td>
<td>2,000</td>
</tr>
</tbody>
</table>

Lambda

- $(\text{errors without knowledge} - \text{errors with knowledge})/\text{errors without knowledge}$
- $(900-300)/900 = .67$
- The .67 indicates that by knowing the respondents’ gender, we have reduced the error in guessing the vote preference by 67%.

Example on .spss
- Spam eating and Vote choice
- 1=eat Spam
- 0=don’t eat spam
- 1=vote Republican
- 0=don’t vote Republican
- Spam= 1, 1, 1, 1, 1, 0, 0, 0, 0, 0,
- Vote= 1, 1, 1, 1, 1, 0, 0, 0, 0, 0,

Another Example on .spss
- Spam= 1, 1, 1, 1, 1, 0, 0, 0, 0, 0,
- Vote= 1, 0, 1, 0, 1, 0, 1, 0, 1, 0,
Wrapping Up

☐ What was the most important point from today?
☐ What would you like to know more about?
☐ What was the muddiest point from today?