

Panel 1

Contingency Tables
Do Males/Females vote differently.

	Male	Female	
Demos	4 $\frac{4}{12}$ $\frac{4}{12}$	3 $\frac{3}{12}$	7
Ind.	4 40% $\frac{4}{12}$	6 60%	10
Repub.	4 $\frac{4}{12}$	2 $\frac{2}{13}$	6
Others	$\frac{2-12}{15} = 0.96$	2	2
	12	13	25

How many of the Demos are male? $\frac{4}{7}$ (Row %)
How many women are indep. $\frac{6}{13}$ (Col %)

1

Panel 2

What is the point of the expected values?

They are those counts that are to be expected if two variables have nothing to do with each other, i.e. if there is no relation between them.

2 Tables: actual counts:

#	#
#	#

 what is actually the case

expected counts:

#	#
#	#

 what are expected

Ex:

10	20
25	10

 actual

11	19
24	11

 expected \Rightarrow No Relation

2

Panel 3

10	20
25	10

actual

5	33
6	19

expected
-> Related

In other words, you compute

$$\chi^2 = \text{sum of } \left(\frac{\text{difference between actual and expected counts}}{\text{expected count}} \right)^2$$

"chi"
 If sum is large => is relation
 If sum is small => no relation

Work is done by Stat Crunch

3

Panel 4

Ex: Gender vs. Pres. Candidate in 2004

Contingency table results:
 Rows: SEX
 Columns: WOULD HAVE VOTED FOR IN 2004

Cell format		Bush	Kerry	Nader	Total
Count	Expected count				
Female	106	170	63	339	
	105.1	166.5	67.37		
Male	89	139	62	290	
	89.9	142.5	57.63		
Total	195	309	125	629	

Chi-Square test:

Statistic	DF	Value	P-value
Chi-square	2	0.78769374	0.6745

If p-value of Chi-Square Test is

$p \geq 0.05 \Rightarrow$ no relation

$p < 0.05 \Rightarrow$ is relation

$p = 0.6745 \Rightarrow$ No relation

Caution: This procedure (Chi-square test) may be invalid if any expected value is less than 5

4

Panel 5

Contingency table results:
 Rows: DEATH PENALTY FOR MURDER
 Columns: LIBERAL OR CONSERVATIVE

Cell format		1 - extremely liberal	2 - liberal	3 - slightly liberal	4 - moderate	5 - slightly conservative	6 - conservative	7 - extremely conservative	Total
favor	Count	28	105	126	468	181	253	53	1214
	Expected count	45.06	151.8	143.8	453.9	169.6	205.4	44.4	
oppose	Count	40	124	91	217	75	57	14	618
	Expected count	22.94	77.25	73.2	231.1	86.36	104.6	22.6	
Total		68	229	217	685	256	310	67	1832

Chi-Square test:

Statistic	DF	Value	P-value
Chi-square	6	109.52265	<0.0001

all expected values ≥ 5 ✓
 $p < 0.0001 \Rightarrow$ There is a relation between variables.

5

Panel 6

Book goes in more detail into this question:

Lambda λ
 Gamma γ values for more precise analysis.

StatCrunch does not compute those λ, γ values, so
 we'll skip them for now.

6

Panel 7

Numeric Variables

Want to know is there a relationship, and if so, make predictions

Ex.	HS GPA	College GPA
Person A:	3.8	2.9
B:	3.1	2.2
C:	4.0	3.5
D:	2.7	1.9
E:	3.3	2.7

a) Are HS + College GPA related?

b) Use HS GPA to predict College GPA

Compute: Correlation Coefficient r
to answer question a)

Create a linear regression model
to answer question b)

7

Panel 8

Def. Correlation Coefficient r .

$$r = \frac{S_{xy}}{\sqrt{S_{xx} \cdot S_{yy}}}$$

$$S_{xx} = \sum x^2 - \frac{(\sum x)^2}{n}$$

$$S_{yy} = \sum y^2 - \frac{(\sum y)^2}{n}$$

$$S_{xy} = \sum xy - \frac{(\sum x)(\sum y)}{n}$$

8