

Panel 1

Last Time

Contingency tables: use variable 1 in rows, var 2 in cols, to find out if they are related!

Row / Col percent

	A	B	C	Totals
1	6			
2	4	7	9	20
2	5			
	15			

$4 \text{ into row \%} = \frac{4}{20} \cdot 100$
 $\text{into col \%} = \frac{4}{15} \cdot 100$

1

Panel 2

	1 - extremely liberal	2 - liberal	3 - slightly liberal	4 - moderate	5 - slightly conservative	6 - conservative	7 - extremely conservative	Total
legal	30 (6.263%) (62.5%)	77 (16.08%) (58.33%)	76 (15.87%) (52.78%)	185 (38.62%) (41.2%)	50 (10.44%) (28.41%)	51 (10.65%) (25%)	10 (2.088%) (25.84%)	479 (100.00%) (40.18%)
not legal	18 (2.525%) (37.5%)	55 (7.714%) (41.67%)	68 (9.537%) (47.22%)	264 (37.03%) (58.8%)	126 (17.87%) (71.59%)	153 (21.46%) (75%)	29 (4.067%) (74.36%)	713 (100.00%) (59.82%)
Total	48 (4.027%) (100.00%)	132 (11.07%) (100.00%)	144 (12.08%) (100.00%)	449 (37.67%) (100.00%)	176 (14.77%) (100.00%)	204 (17.11%) (100.00%)	39 (3.272%) (100.00%)	1192 (100.00%) (100.00%)

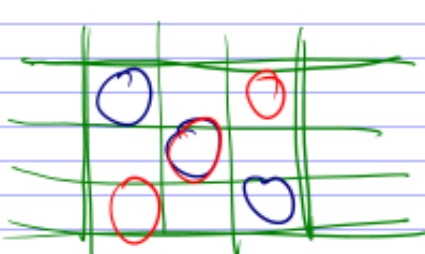
Should marijuana be legalized

- ① Which number is row %, col %
- ② What is the missing number?
- ③ Of all the supporters of legalizing marijuana, how many are moderates? ~~185~~ 38.62% (row %)
- ④ How many liberals are against legalizing? ~~55~~ (41.67%)

Panel 3

Big Question: Is there a relationship between two variables.

Do this:
 indep var \Rightarrow col
 dep var \Rightarrow row
 compute row %
 circle highest value per row



direct relation between them
 inverse relationships between them.

3

Panel 4

Does Money make you happy?

\uparrow
indep
 \uparrow
dep.

Rows: SATISFACTION WITH FINANCIAL SITUATION
 Columns: GENERAL HAPPINESS

Cell format					
Count (Row percent)					
	1 - very happy	2 - pretty happy	3 - not too happy	Total	
1 - satisfied	267 (46.76%)	273 (47.81%)	31 (5.429%)	571 (100.00%)	0
2 - more or less	236 (28.99%)	476 (58.48%)	102 (12.53%)	814 (100.00%)	
3 - not at all satisfied	93 (14.93%)	348 (55.86%)	182 (29.21%)	623 (100.00%)	
Total	596 (29.88%)	1097 (54.63%)	315 (15.69%)	2008 (100.00%)	

not related!

4

Panel 5

Does your political opinion impact your view on legalizing Marijuana?

	1 - extremely liberal	2 - liberal	3 - slightly liberal	4 - moderate	5 - slightly conservative	6 - conservative	7 - extremely conservative	Total
legal	70 (6.268%)	77 (16.08%)	76 (15.87%)	189 (38.62%)	50 (10.44%)	51 (10.65%)	10 (2.088%)	479 (100.00%)
not legal	18 (2.525%)	65 (7.714%)	68 (9.537%)	264 (37.03%)	126 (17.67%)	159 (21.46%)	24 (4.067%)	713 (100.00%)
Total	48 (4.027%)	132 (11.07%)	144 (12.08%)	449 (37.67%)	176 (14.77%)	204 (17.11%)	39 (3.272%)	1192 (100.00%)

5

Panel 6

This does not work super well, especially only for ordinal vars.

Contingency mechanism \Rightarrow expected values

	female	males	
Smoke	$\frac{40}{100} \cdot 30$	$\frac{60}{100} \cdot 30$	30
not smoke	$\frac{40}{100} \cdot 70$	$\frac{60}{100} \cdot 70$	70
	40	60	100

If there was no relation between row, col variables \Rightarrow expected value

$$\frac{(\text{row total}) \times (\text{col total})}{\text{total}}$$

6

Panel 7

Contingency Tables
Do Males/Females vote differently.

	Male	Female	
Demos	4 $\frac{7 \cdot 12}{25}$	3 $\frac{7 \cdot 13}{25}$	7
Ind.	4	6	10
Repub.	4 $\frac{6 \cdot 12}{25}$	2	6
Others	2 $\frac{2 \cdot 12}{25}$	2 $\frac{2 \cdot 13}{25}$	2
	12	13	25

How many of the Demos are male? $\frac{4}{7}$ row %

How many women are indep? $\frac{6}{13}$ col %

Find some expected values?
 $\frac{(\text{row total})(\text{col total})}{\text{total}}$

Panel 8

What is the point of the expected values?

Generate 2 tables:

#	#
#	#

actual numbers as found in data

#	#
#	#

expected values, i.e. values I expect to see if there is no relation!

Ex:

10	20
25	10

 actual

11	19
24	11

 no relation

Panel 9

10	20
25	10

actual

5	33
6	19

expected

yes, related

chi-squared

In other words, compute

$$\text{sum of (rel. diff. between actual and expected value)}^2 = \chi^2$$

if χ^2 is large \Rightarrow yes, related

if χ^2 is small \Rightarrow no, related

\Rightarrow Statcrunch does this easily!

Panel 10

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	105.1	106	170	63	339
Male	89.9	139	142.5	57.5	339
Total		195	309	125	629

Chi-Square test:

Statistic	DF	Value	P-value
Chi-square	2	0.76769374	0.6745

If p-value of Chi-Square Test is

$p \geq 0.05 \Rightarrow (\chi^2 \text{ is small}) \Rightarrow$ no relation

$p < 0.05 \Rightarrow (\chi^2 \text{ is large}) \Rightarrow$ they are related

$105.1 = \frac{339 \cdot 195}{629}$

no relation

Panel 11

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

Cell format								
Count	Expected count							
	1 - extremely liberal	2 - liberal	3 - slightly liberal	4 - moderate	5 - slightly conservative	6 - conservative	7 - extremely conservative	Total
favor	28 45.06	105 151.8	126 143.8	468 453.9	181 169.6	253 205.4	53 44.4	1214
oppose	40 22.94	124 77.25	91 73.2	217 231.1	75 86.38	57 104.6	14 22.6	618
Total	68	229	217	685	256	310	67	1832

Chi-Square test:

Statistic	DF	Value	P-value
Chi-square	6	109.5225	0.0001

They are related

11

Panel 12

12