CROSSTABS – CHI^2

And manipulating the data

See my book for an example.

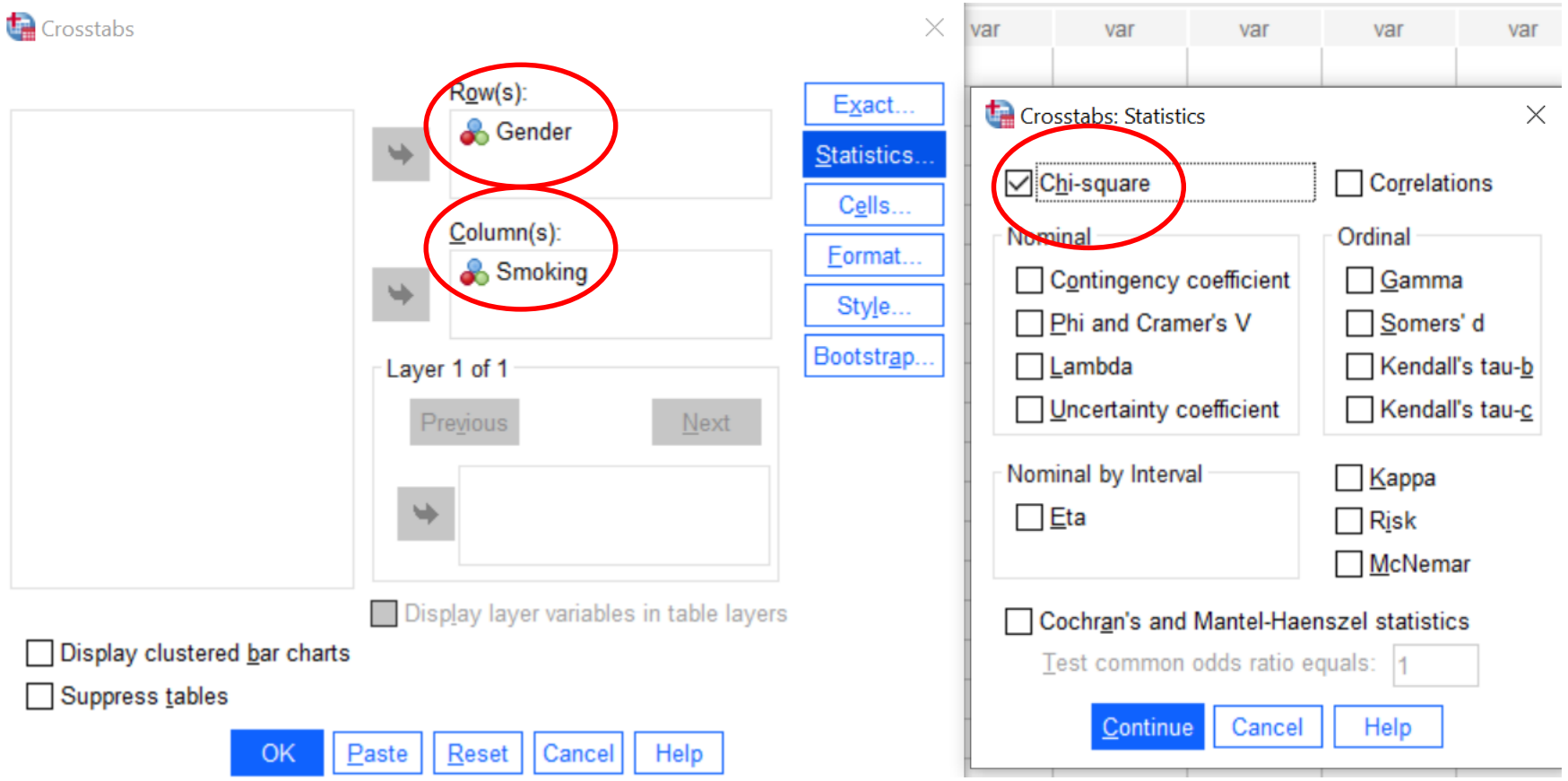
Research question

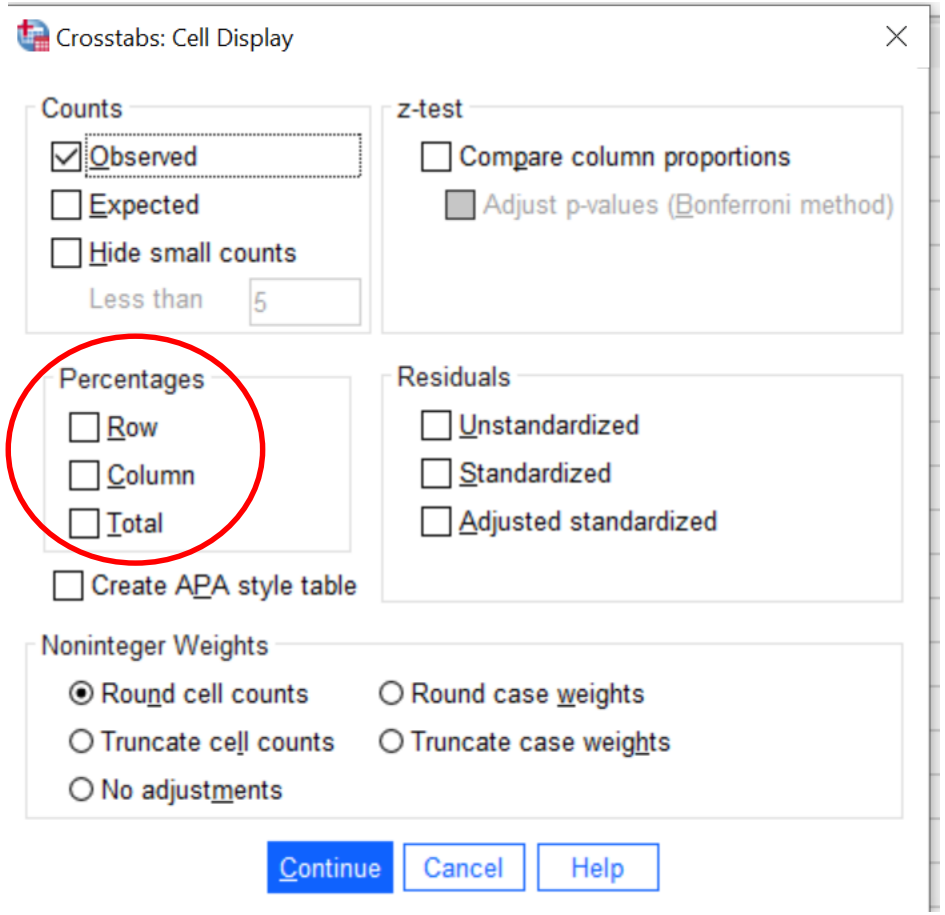
- Is there a difference between genders for smoking?

Crosstabs

- Two categorical (nominal or ordinal) variables
- Try to stick with 2x2 tables.
- Larger tables are possible, but explanation becomes quickly complex.

- Our example:
  - Men versus women
  - No smoke versus smoke


Crosstabs
Row(s):
Gender
Column(s):
Smoking
Layer 1 of 1
Previous
Next
Display layer variables in table layers
Display clustered bar charts
Suppress tables
OK
Paste
Reset
Cancel
Help
Exact...
Statistics...
Cells...
Format...
Style...
Bootstrap...
var
Crosstabs: Statistics
Chi-square
Correlations
Nominal
Contingency coefficient
Phi and Cramer's V
Lambda
Uncertainty coefficient
Ordinal
Gamma
Somers' d
Kendall's tau-b
Kendall's tau-c
Nominal by Interval
Eta
Kappa
Risk
McNemar
Cochran's and Mantel-Haenszel statistics
Test common odds ratio equals:
1
Continue
Cancel
Help


Crosstabs: Cell Display
Counts
Observed
Expected
Hide small counts
Less than
5
z-test
Compare column proportions
Adjust p-values (Bonferroni method)
Percentages
Row
Column
Total
Residuals
Unstandardized
Standardized
Adjusted standardized
Create APA style table
Noninteger Weights
Round cell counts
Round case weights
Truncate cell counts
Truncate case weights
No adjustments
Continue
Cancel
Help

Assumption for 2x2: At least (5)10 observations in each cell.

Perhaps too few observations

Gender * Smoking Crosstabulation

			Smoking		
			Some	No Smoke	Total
Gender	men	Count	34	32	66
		% within Gender	51,5%	48,5%	100,0%
		% within Smoking	85,0%	53,3%	66,0%
		% of Total	34,0%	32,0%	66,0%
	Women	Count	6	28	34
		% within Gender	17,6%	82,4%	100,0%
		% within Smoking	15,0%	46,7%	34,0%
		% of Total	6,0%	28,0%	34,0%
Total		Count	40	60	100
		% within Gender	40,0%	60,0%	100,0%
		% within Smoking	100,0%	100,0%	100,0%
		% of Total	40,0%	60,0%	100,0%

There is too much information to clearly interpret the results.

2x2 matrix so continuity correction

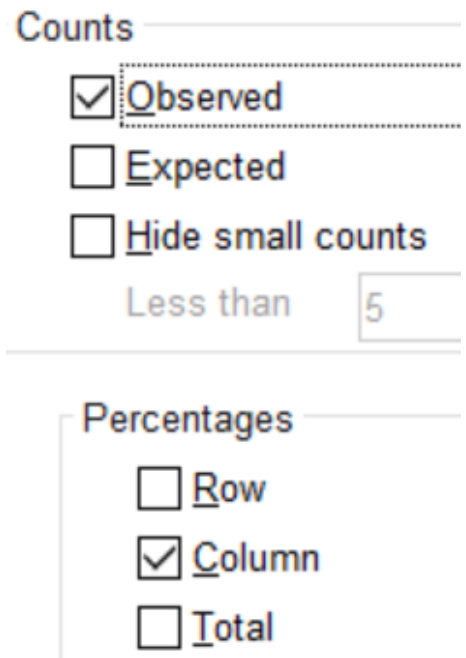
Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	10,725[a]	1	,001		
Continuity Correction[b]	9,360	1	,002		
Likelihood Ratio	11,480	1	<,001		
Fisher's Exact Test				,001	<,001
Linear-by-Linear Association	10,618	1	,001		
N of Valid Cases	100				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 13,60.

b. Computed only for a 2x2 table

Conclusion: .002<.05, so there is a statistically significant difference between genders for smoking. Men smoke more than women.



Gender * Smoking Crosstabulation

			Smoking		
			Smoke	No Smoke	Total
Gender	men	Count	34	32	66
		% within Smoking	85,0%	53,3%	66,0%
	Women	Count	6	28	34
		% within Smoking	15,0%	46,7%	34,0%
Total		Count	40	60	100
		% within Smoking	100,0%	100,0%	100,0%

Of smokers, 85% are male, whereas for non-smokers they are about evenly disributed across gender (53.3% male and 46.7% female).

Percentages

- [x] Row
- [ ] Column
- [ ] Total

Gender * Smoking Crosstabulation

			Smoking		
			Smoke	No Smoke	Total
Gender	men	Count	34	32	66
		% within Gender	51,5%	48,5%	100,0%
	Women	Count	6	28	34
		% within Gender	17,6%	82,4%	100,0%
Total		Count	40	60	100
		% within Gender	40,0%	60,0%	100,0%

Men are about equally divided between smokers (51.5%) and non-smokers (48.5%). For women, substantially fewer are smokers (17.6%) than non-smokers (82.4%).

Percentages

- [ ] Row
- [ ] Column
- [x] Total

Gender * Smoking Crosstabulation

			Smoking		
			Smoke	No Smoke	Total
Gender	men	Count	34	32	66
		% of Total	34,0%	32,0%	66,0%
	Women	Count	6	28	34
		% of Total	6,0%	28,0%	34,0%
Total		Count	40	60	100
		% of Total	40,0%	60,0%	100,0%

In the sample, there are fewer smokers (40%) than non-smokers (60%), and the sample has more men (66%) than women (34%).