

Making Contingency Tables in Stata

On contingency tables, see Agresti/Finlay, *Statistical Methods for the Social Sciences*, chap. 8; and Ian Watson, "Publications quality tables in Stata: a tutorial for the *tabout* program"

http://www.ianwatson.com.au/stata/about_tutorial.pdf

```
use hsb2, clear
tab1 ses prog
```

tabulate

```
tab prog ses, col all
```

This set of options uses column format and displays Pearson's chi-squared, likelihood-ratio, Cramer's V, gamma, and Kendall's tau-b (see Agresti/Finlay).

tabi for 'immediate' form

Options:

col, row, cell: results format
nofreq: do not display frequencies
chi2: Pearson chi-squared test
exact: Fisher's exact test
gamma: gamma test
taub: Kendall's tau-b test
V: Cramer's V test

tabchi

```
findit tabchi
help tabchi
```

```
tabchi prog ses, adjust
```

tabchi displays expected frequencies, observed frequencies, and adjusted frequencies, as well as chi-squared and likelihood-ratio tests of significance (see Agresti/Finlay).

tabchii for 'immediate' form.

Options:

raw: display raw residuals
pearson: display Pearson residuals
adjust: display adjusted residuals
noe: do not display expected frequencies
noo: do not display observed frequencies

tabout

To make publication-style tables. See the excellent *tabout* tutorial:
http://www.ianwatson.com.au/stata/about_tutorial.pdf

findit about

help about

la var female "Gender"

la var prog "Program"

tabout female prog using gender_prog, cell(freq col) format(0 1) stats(chi2)
layout(row) show(all) replace

Edit file 'gender_prog' (located in current folder) in Excel or Word, or generate LaTeX output. The following works very well: In Excel, File>Open; select the about file from the current folder, and use Excel's wizard to format the table in Excel. Either directly edit or use the Excel table, or copy the table in Excel and then 'paste special' into Word.

Some options:

cell: freq, cell, row, col, cum [all can be specified]

format: # decimal points; specify in order of types of cell entries specified

clab: table's title

layout: how the columns are laid out stats: types of statistics specified

The following about guidelines are taken from Ian Watson, "Publications quality tables in Stata: a tutorial for the about program"

http://www.ianwatson.com.au/stata/about_tutorial.pdf

Type of table	Allowable cell contents	Available layout
Basic	freq cell row col cum <i>Any number of these, in any order</i>	col row cb rb
Basic with SE or CI (turn on svy option)	freq cell row col se ci lb up <i>Only one of: freq cell row col (any of which must come first in the cell) And any number of: se ci lb up E.g.: cells(col se lb up)</i>	col row cb rb
Summary (oneway table)	<i>Any number of: N mean var sd sum min max count median iqr p25 p50 p75 (and others)</i>	no options (fixed)

(Continued on next page)

(Continued from previous page)

Type of table	Allowable cell contents	Available layout
	<i>With each followed by variable name E.g.: cells(min wage mean age)</i>	
Summary	<i>Only one of: N mean var sd sum min max count median iqr p25 p25 p50 p75 (and others) Followed by one variable name: E.g.: cells(sum income)</i>	no options (fixed)
Summary with	<i>mean followed by one variable and any number of: se ci lb ub E.g.: cells(mean weight se ci)</i>	col row cb rb

See Watsons' extensive discussion and examples.

tabstat

tabstat displays summary statistics for a series of numeric variables in a single table, possibly broken down on (conditioned by) another variable.

help tabstat

tabstat math, stats(N mean sd median min max) by(ses) format(%9.2f)

Some options:

nototal missing column(stats) column(variables)

fsum

fsum read write math science, stats(mean med sd min max) format(%9.2f)

bys female: fsum read write math science, stats(mean med sd min max) format(%9.2f)

statsmat

```
statsmat read write math science, mat(matrix1) s(mean sd median p25 p75 min  
max) f(%3.2f)
```

makematrix

```
makematrix, from(r(mean) r(p50) r(sd) r(min) r(max)) format(%9.2f): su read-  
science, detail
```

```
makematrix, from(r(rho)): corr read-socst
```

```
makematrix, from(r(rho)) cols(socst) format(%9.2f): corr read-socst
```