PUBH 501 Biostatistics

STATA: CHI-SQUARE, FISHER'S EXACT, AND MCNEMAR'S TESTS

Tip of the day

- Copy and paste from Stata by highlighting output from the results window.
- Right click and select “copy as table.” This will allow you to paste into Excel as a table, instead of plain text. See below…

Education: |			
0 if HS/GED |			
or more, 1 |			
if did not |			
finish HS |	Freq.	Percent	Cum.
----------+----------			
0 |	209	59.54	59.54
1 |	142	40.46	100.00
----------+----------			
Total |	351	100.00	

Education:			
0 if HS/GED			
or more, 1			
if did not			
finish HS	Freq.	Percent	Cum.
0	209	59.54	59.54
1	142	40.46	100
Total	351	100	

Overview

- Chi-square tests
  - Goodness of fit
  - Contingency
- Fisher's exact test
- McNemar's test

Chi-square assumptions

- Categorical variable: binary, nominal, or ordinal
- Mutually exclusive categories
- Independent sample
- The smallest number in an expected cell is > 5

Checking assumptions in Stata

- Use the –tabulate- command to look at the variable(s) of interest

Goodness of fit command

- There is no native command in Stata for Chi-square goodness of fit test
- Install the user-written command –csgof-
  - findit csgof
  - Click the link that says “csgof from https://stats.idre.ucla.edu/stat/stata/ado/analysis”
  - Another window will pop up. Click the link that says click here to install.

Viewer - search csgof, all

File Edit History Help

search csgof, all

search csgof, all

search for **csgof**

Search of official help files, FAQs, Examples, and Stata Journals

FAQ . . . . . What statistical analysis should I use?
. . . . . UCLA Academic Technology Services
5/08 https://stats.idre.ucla.edu/stata/whatstat/what-statistical-analysis-should-i-usestatistical-analyses-using-stata/

Search of web resources from Stata and other users

(contacting http://www.stata.com)

1 package found (Stata Journal and STB listed first)

csgof from https://stats.idre.ucla.edu/stat/stata/ado/analysis
csgof. Chi-Square Goodness of Fit / Statistical Consulting Group /
Institute for Digital Research and Education, UCLA / idrestat@ucla.edu /
STATA ado and hlp files in the package / distribution-date: 20150326

(click here to return to the previous screen)

(end of search)

Viewer - net describe csgof, from(https://stats.idre.ucla.edu/stat/stata/ado/analysis)

File Edit History Help

net describe csgof, from(https://stats.idre.ucla.edu/stat/stata/ado/analysis)

net describe csgof, from(https:/...

---

package **csgof** from https://stats.idre.ucla.edu/stat/stata/ado/analysis

---

TITLE

csgof. Chi-Square Goodness of Fit

DESCRIPTION/AUTHOR(S)

Statistical Consulting Group
Institute for Digital Research and Education, UCLA
idrestat@ucla.edu

STATA ado and hlp files in the package
distribution-date: 20150326

INSTALLATION FILES (click here to install)

csgof.ado
csgof.hlp

(click here to return to the previous screen)

Goodness of fit command

- The user (you) must enter the expected frequencies. In the case below, they are 70, 20, and 10%

- `csgof race, expperc(70, 20, 10)`

```
  +-------------------------------------+
  |  race   expperc   expfreq   obsfreq |
  |-------------------------------------|
  | white        70     132.3         96 |
  | black        20      37.8         26 |
  | other        10      18.9         67 |
  +-------------------------------------+

chisq(2) is 136.06, p = 0
```

Chi-square contingency / test of independence

- Two-sample test: compare proportions between two groups
- Use the –chi2- option in the –tabulate- command (remember, -tab- for short)
- Can also use the –row- and –column- options to show the row and column percentages

- `tab ui low, row col chi2`

Results

Uterine irritability	birthweight<2500g		
no	116	45	161
	72.05	27.95	100.00
	89.23	76.27	85.19
yes	14	14	28
	50.00	50.00	100.00
	10.77	23.73	14.81
Total	130	59	189
	68.78	31.22	100.00
	100.00	100.00	100.00

Pearson chi2(1) = 5.4008 Pr = 0.020

Fisher’s exact test

Fisher’s exact test

- Used in place of chi-square where expected frequency of any cell is < 5
- Use the –expected- option in the –tab- command to look at expected cell size
- Use the –exact- option in the –tab- command to run the Fisher’s exact test

- `tab ui low, expected`
- `tab ui low, exact`

Results

hypertension	birthweight<2500g		
	no	yes	Total
no	125	52	177
	121.7	55.3	177.0
yes	5	7	12
	8.3	3.7	12.0
Total	130	59	189
	130.0	59.0	189.0

Fisher's exact = 0.052

1-sided Fisher's exact = 0.042

McNemar’s test

McNemar’s test

- Test change in proportion for matched or paired data
- Variables must be coded as 0/1
- Look at variables using –tab- command.
- The –mcc- command is used for the test. mcc stands for matched case control

- `mcc cesd0dich23 cesd1dich23`

Results

Cases	Controls: Exposed	Controls: Unexposed	Total
Exposed	3	8	11
Unexposed	3	36	39
Total	6	44	50

McNemar's chi2(1) = 2.27 Prob > chi2 = 0.1317
Exact McNemar significance probability = 0.2266

~~Proportion with factor~~
~~Cases .22~~
~~Controls .12 [95% Conf. Interval]~~
~~difference .1 -.0470202 .2470202~~
~~ratio 1.833333 .8236446 4.080778~~
~~rel. diff. .1136364 -.0254542 .2527269~~

~~odds ratio 2.666667 .6400364 15.6064 (exact)~~

Use the "exact" *p*-value when the sum of the discordant pairs (exposed case and unexposed control, and vice versa) is < 20

Ignore crossed out text for now. This will come up in the Stata output.