Subject: How to make and interpreting Bivariate statistics analysis?
Posted by hamzah on Mon, 12 Dec 2016 19:14:49 GMT

Dear experts

Regarding statistics to population survey, could you please tell me which one of the syntax using for bivariate analysis [chi square] and what does different the meaning of each syntax like below:

1.
svy: tabulate sex malaria
and output here :

```
Number of strata  =      1            Number of obs   =   259,885
Number of PSUs    =   4,418           Population size = 30,152,652
                            Design df      =    4,417
```

```
------------------------------
gender of |
responden |     malaria
ts     |  no   yes  Total
----------+--------------------
   male | .4744  .0185  .4929
 female | .4909  .0162  .5071
        |
  Total | .9653  .0347     1
------------------------------
  Key:  cell proportion

  Pearson:
    Uncorrected   chi2(1)     =   58.3020
    Design-based  F(1, 4417)    =   49.6352    P = 0.0000
```

2.
. svy: tabulate sex malaria, row
and output here :
(running tabulate on estimation sample)

```
Number of strata  =      1            Number of obs   =   259,885
Number of PSUs    =   4,418           Population size = 30,152,652
                            Design df      =    4,417
```

```
------------------------------
gender of |
responden |     malaria
ts     |  no   yes  Total
```

```
----------+--------------------
  male | .9625  .0375    1
female | .968   .032     1
       |
 Total | .9653  .0347    1
------------------------------
Key:  row proportion


Pearson:
  Uncorrected  chi2(1)      =  58.3020
  Design-based  F(1, 4417)   =  49.6352    P = 0.0000
```

3.
. svy linearized : tabulate sex malaria, obs row percent ci

and output here :
 (running tabulate on estimation sample)

```
Number of strata  =      1          Number of obs    =    259,885
Number of PSUs    =   4,418           Population size  = 30,152,652
                          Design df       =    4,417
```

```
---------------------------------------------------------
gender of |
responden |              malaria
ts    |        no        yes      Total
----------+----------------------------------------------
  male |       96.25       3.746         100
       | [96.01,96.48] [3.518,3.987]
       |       1.2e+05        5595     1.3e+05
       |
female |       96.8        3.198         100
       | [96.57,97.02] [2.979,3.431]
       |       1.3e+05        4971     1.3e+05
       |
 Total |       96.53       3.468         100
       | [96.31,96.74] [3.257,3.692]
       |       2.5e+05      1.1e+04     2.6e+05
---------------------------------------------------------
Key:  row percentage
      [95% confidence interval for row percentage]
      number of observations


Pearson:
```

```
Uncorrected  chi2(1)     =   58.3020
Design-based  F(1, 4417)    =  49.6352    P = 0.0000
```

How to make odds ratio for cross-sectional design survey? Should I make syntax for prevalence ratio or may I take directly odds ratio in the syntax below?

5.
. svy linearized : logistic sex malaria

and output here :
 (running logistic on estimation sample)

Survey: Logistic regression

```
Number of strata  =      1              Number of obs    =     259,885
Number of PSUs    =   4,418              Population size  = 30,152,652
                            Design df        =     4,417
                            F(  1,  4417)  =     49.54
                            Prob > F        =   0.0000


------------------------------------------------------------------------------
            |           Linearized
      sex  | Odds Ratio  Std. Err.    t   P>|t|    [95% Conf. Interval]
-------------+----------------------------------------------------------------
   malaria |  .8488294  .0197667   -7.04  0.000    .8109481    .8884803
     _cons |  1.034818  .0042681    8.30  0.000    1.026484    1.043219
------------------------------------------------------------------------------
```

Based on the table above [chi square and binary logistic].

Where the sex variable which assumptions male is given code = 0 and female is given code = 1. Malaria prevalence differs by sex Males are more likely to have malaria than females (1.85% males versus 1.62% females, P = 0.000). Based on odds ratio (OR) female have the chances of getting malaria 0.85% or 0.85 times than male (as categorical reference)

How do I interpret an odds ratio less than 1 in a logistic regression?
May I will be written male with a chance of 1 / 0.85 times or 1.2 times to get malaria compared than female as well?

or

The odds of malaria in male decreased by (1 - 0.85) 15% compared those in a female. Whatever on the dependent variable decreases. For each unit increase, it decreases by a multiple of (1 -

OR)

Thank you in advance for your reply

Sincerely yours,

Hamzah

Dear User,

We recommend that you search online for some resources to better understand the output you are generating. This site is a good reference with lots of examples and help with interpretation: http://statistics.ats.ucla.edu/stat/stata/

A few notes on your output:
1) Without more information on the dataset you are using and on your analytic objectives it is difficult to provide much guidance but it appears that your sample size is unusually large. Is this a pooled sample from multiple surveys?
2) It also appears that you have misrepresented the sample design in your survey set command. It would be very unusual for a DHS or MIS survey to have only one stratum.
3) I would also suggest verifying that you are properly specifying the denominator for this analysis. In most DHS/MIS surveys, only children 6-59 months of age who spent the previous night in interviewed households (de facto) were eligible for malaria parasitemia testing.
4) Please check that you have properly recoded the sex variable. In our standard recode files 1=male and 2=female.
5) For proper interpretation of ORs <1.0 please see online resources such as the one given above.

Good luck!