svy: tabulate oneway — One-way tables for survey data

Syntax

Basic syntax

    svy: tabulate varname

Full syntax

    svy [vcetype] [, svy_options] : tabulate varname [if] [in]
    [ , tabulate_options display_items display_options]

Syntax to report results

    svy [, display_items display_options]

description

vcetype                Description

SE

linearized           Taylor-linearized variance estimation
bootstrap            bootstrap variance estimation; see [SVY] svy bootstrap
brr                    BRR variance estimation; see [SVY] svy brr
jackknife             jackknife variance estimation; see [SVY] svy jackknife
sdr                   SDR variance estimation; see [SVY] svy sdr

Specifying a vcetype overrides the default from svyset.
### svy_options

<table>
<thead>
<tr>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>if/in</strong></td>
</tr>
<tr>
<td><strong>subpop([varname] [if])</strong></td>
</tr>
</tbody>
</table>

#### SE

<table>
<thead>
<tr>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>bootstrap_options</strong></td>
</tr>
<tr>
<td><strong>brr_options</strong></td>
</tr>
<tr>
<td><strong>jackknife_options</strong></td>
</tr>
<tr>
<td><strong>sdr_options</strong></td>
</tr>
</tbody>
</table>

svy requires that the survey design variables be identified using `svyset`; see [SVY] svyset.

See [U] 20 Estimation and postestimation commands for more capabilities of estimation commands.

Warning: Using if or in restrictions will often not produce correct variance estimates for subpopulations. To compute estimates for subpopulations, use the `subpop()` option.

### tabulate_options

<table>
<thead>
<tr>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model</strong></td>
</tr>
<tr>
<td><strong>stdize(varname)</strong></td>
</tr>
<tr>
<td><strong>stdweight(varname)</strong></td>
</tr>
<tr>
<td><strong>tab(varname)</strong></td>
</tr>
<tr>
<td><strong>missing</strong></td>
</tr>
</tbody>
</table>

### display_items

<table>
<thead>
<tr>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Table items</strong></td>
</tr>
<tr>
<td><strong>cell</strong></td>
</tr>
<tr>
<td><strong>count</strong></td>
</tr>
<tr>
<td><strong>se</strong></td>
</tr>
<tr>
<td><strong>ci</strong></td>
</tr>
<tr>
<td><strong>deff</strong></td>
</tr>
<tr>
<td><strong>deft</strong></td>
</tr>
<tr>
<td><strong>cv</strong></td>
</tr>
<tr>
<td><strong>srssubpop</strong></td>
</tr>
<tr>
<td><strong>obs</strong></td>
</tr>
</tbody>
</table>

When any of se, ci, deff, deft, cv, or srssubpop is specified, only one of cell or count can be specified. If none of se, ci, deff, deft, cv, or srssubpop is specified, both cell and count can be specified.
svy: tabulate oneway — One-way tables for survey data

<table>
<thead>
<tr>
<th>display_options</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>level(#)</td>
<td>set confidence level; default is level(95)</td>
</tr>
<tr>
<td>proportion</td>
<td>display proportions; the default</td>
</tr>
<tr>
<td>percent</td>
<td>display percentages instead of proportions</td>
</tr>
<tr>
<td>nomarginal</td>
<td>suppress column marginal</td>
</tr>
<tr>
<td>nolabel</td>
<td>suppress displaying value labels</td>
</tr>
<tr>
<td>cellwidth(#)</td>
<td>cell width</td>
</tr>
<tr>
<td>csepwidth(#)</td>
<td>column-separation width</td>
</tr>
<tr>
<td>stubwidth(#)</td>
<td>stub width</td>
</tr>
<tr>
<td>format(%)</td>
<td>cell format; default is format(%6.0g)</td>
</tr>
</tbody>
</table>

proportion is not shown in the dialog box.

Menu

Statistics > Survey data analysis > Tables > One-way tables

Description


Options

svy_options; see [SVY] svy.

stdize(varname) specifies that the point estimates be adjusted by direct standardization across the strata identified by varname. This option requires the stdweight() option.

stdweight(varname) specifies the weight variable associated with the standard strata identified in the stdize() option. The standardization weights must be constant within the standard strata.

tab(varname) specifies that counts be cell totals of this variable and that proportions (or percentages) be relative to (that is, weighted by) this variable. For example, if this variable denotes income, then the cell “counts” are instead totals of income for each cell, and the cell proportions are proportions of income for each cell.

missing specifies that missing values in varname be treated as another row category rather than be omitted from the analysis (the default).

cell requests that cell proportions (or percentages) be displayed. This is the default if count is not specified.

count requests that weighted cell counts be displayed.

se requests that the standard errors of cell proportions (the default) or weighted counts be displayed. When se (or ci, deff, deft, or cv) is specified, only one of cell or count can be selected. The standard error computed is the standard error of the one selected.
ci requests confidence intervals for cell proportions or weighted counts.

deff and deft request that the design-effect measures DEFF and DEFT be displayed for each cell proportion or weighted count. See [SVY] estat for details.

The deff and deft options are not allowed with estimation results that used direct standardization or poststratification.

cv requests that the coefficient of variation be displayed for each cell proportion, count, or row or column proportion. See [SVY] estat for details.

srssubpop requests that DEFF and DEFT be computed using an estimate of SRS (simple random sampling) variance for sampling within a subpopulation. By default, DEFF and DEFT are computed using an estimate of the SRS variance for sampling from the entire population. Typically, srssubpop would be given when computing subpopulation estimates by strata or by groups of strata.

obs requests that the number of observations for each cell be displayed.

level(#) specifies the confidence level, as a percentage, for confidence intervals. The default is level(95) or as set by set level; see [U] 20.7 Specifying the width of confidence intervals.

proportion, the default, requests that proportions be displayed.

percent requests that percentages be displayed instead of proportions.

nomarginal requests that the column marginal not be displayed.

nolabel requests that variable labels and value labels be ignored.

cellwidth(#), csepwidth(#), and stubwidth(#) specify widths of table elements in the output; see [P] tabdisp. Acceptable values for the stubwidth() option range from 4 to 32.

format(%fmt) specifies a format for the items in the table. The default is format(%6.0g). See [U] 12.5 Formats: Controlling how data are displayed.

svy: tabulate uses the tabdisp command (see [P] tabdisp) to produce the table. Only five items can be displayed in the table at one time. The ci option implies two items. If too many items are selected, a warning will appear immediately. To view more items, redisplay the table while specifying different options.

Remarks and examples

Despite the long list of options for svy: tabulate, it is a simple command to use. Using the svy: tabulate command is just like using tabulate to produce one-way tables for ordinary data. The main difference is that svy: tabulate computes standard errors appropriate for complex survey data.

Standard errors and confidence intervals can optionally be displayed for weighted counts or cell proportions. The confidence intervals for proportions are constructed using a logit transform so that their endpoints always lie between 0 and 1; see [SVY] svy: tabulate twoway. Associated design effects (DEFF and DEFT) can be viewed for the variance estimates.

Example 1

Here we use svy: tabulate to estimate the distribution of the race category variable from our NHANES II dataset (McDowell et al. 1981). Before calling svy: tabulate, we use svyset to declare the survey structure of the data.
. use http://www.stata-press.com/data/r13/nhanes2b
. svyset psuid [pweight=finalwgt], strata(stratid)

  pweight: finalwgt
  VCE: linearized
  Single unit: missing
  Strata 1: stratid
  SU 1: psuid
  FPC 1: <zero>

. svy: tabulate race
(running tabulate on estimation sample)

Number of strata = 31  Number of obs      = 10351
Number of PSUs   = 62   Population size  = 117157513
Design df        = 31

1=white,
2=black,
3=other

<table>
<thead>
<tr>
<th></th>
<th>proportions</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>.8792</td>
</tr>
<tr>
<td>Black</td>
<td>.0955</td>
</tr>
<tr>
<td>Other</td>
<td>.0253</td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
</tr>
</tbody>
</table>

Key: proportions = cell proportions

Here we display weighted counts for each category of race along with the 95% confidence bounds, as well as the design effects DEFF and DEFT. We also use the format() option to improve the look of the table.

. svy: tabulate race, format(%11.3g) count ci deff deft
(running tabulate on estimation sample)

Number of strata = 31  Number of obs      = 10351
Number of PSUs   = 62   Population size  = 117157513
Design df        = 31

<table>
<thead>
<tr>
<th></th>
<th>count</th>
<th>lb</th>
<th>ub</th>
<th>deff</th>
<th>deft</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>102999549</td>
<td>97060400</td>
<td>108938698</td>
<td>60.2</td>
<td>7.76</td>
</tr>
<tr>
<td>Black</td>
<td>11189236</td>
<td>8213964</td>
<td>14164508</td>
<td>18.6</td>
<td>4.31</td>
</tr>
<tr>
<td>Other</td>
<td>2968728</td>
<td>414930</td>
<td>5522526</td>
<td>47.9</td>
<td>6.92</td>
</tr>
<tr>
<td>Total</td>
<td>117157513</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Key: count = weighted counts
lb = lower 95% confidence bounds for weighted counts
ub = upper 95% confidence bounds for weighted counts
deff = deff for variances of weighted counts
deft = deft for variances of weighted counts

From the above results, we can conclude with 95% confidence that the number of people in the population that fall within the White category is between 97,060,400 and 108,938,698.
Stored results

In addition to the results documented in [SVY] svy, svy: tabulate stores the following in e():

Scalars
- e(r) number of rows
- e(total) weighted sum of tab() variable

Macros
- e(cmd) tabulate
- e(tab) tab() variable
- e(rowlab) label or empty
- e(rowvlab) row variable label
- e(rowvar) varname, the row variable
- e(setype) cell or count

Matrices
- e(Prop) matrix of cell proportions
- e(Obs) matrix of observation counts
- e(Deff) DEFF vector for e(setype) items
- e(Deft) DEFT vector for e(setype) items
- e(Row) values for row variable
- e(V_row) variance for row totals
- e(V_srs_row) V_srs for row totals
- e(Deff_row) DEFF for row totals
- e(Deft_row) DEFT for row totals

Methods and formulas

See Methods and formulas in [SVY] svy: tabulate twoway for a discussion of how table items and confidence intervals are computed. A one-way table is really just a two-way table that has one row or column.

Reference


Also see

[SVY] svy postestimation — Postestimation tools for svy
[SVY] svydescribe — Describe survey data
[R] tabulate oneway — One-way table of frequencies
[SVY] svy: tabulate twoway — Two-way tables for survey data
[U] 20 Estimation and postestimation commands
[SVY] direct standardization — Direct standardization of means, proportions, and ratios
[SVY] poststratification — Poststratification for survey data
[SVY] subpopulation estimation — Subpopulation estimation for survey data
[SVY] svy — The survey prefix command
[SVY] variance estimation — Variance estimation for survey data