Subject: svy command and descriptives
Posted by shayankhan on Sat, 08 Oct 2016 20:35:29 GMT
View Forum Message <> Reply to Message

I know we have to use the svy command to get the descriptives. I am studying children under 5 years of age of Pakistan for ARI. Now, my criteria for ARI differs from the one used in the PDHS report. I have used the criteria in the report and my results match that way. Anyway, my question is that when I run the svy: tab command or tab with weights command, I get different population sizes for different explanatory variables. The population size for children under 5 years of age is 11,040. So, I want to construct my descriptives table according to that sample but I cannot as the percentages than considers the population size for each explanatory variable respectively. How do I construct my table? Please help.

Bridgette, please help me out. I am in a hurry. This is my thesis.

Subject: Re: svy command and descriptives
Posted by Bridgette-DHS on Mon, 10 Oct 2016 11:57:05 GMT
View Forum Message <> Reply to Message

Following is a response from Senior DHS Stata Specialist, Tom Pullum:

Quote:When you combine two or more tables, variables that are "." (not applicable) on any of the variables will be excluded. The KR file, which I believe you are using, includes all children born in the past five years, whether they are alive or not. That is probably the 11,040 you are talking about (I don't have time to check that). The questions on ARI, etc., are limited to children who are alive. Some questions in the KR file are limited further to children who are alive and living with the mother, or limited to the youngest living child. Children for whom the question does not apply are coded ".". That's why you are getting different n's. That has nothing to do with svy. You could repeat without svy and you would see the same thing.

Subject: Re: svy command and descriptives
Posted by shayankhan on Mon, 10 Oct 2016 12:12:35 GMT
View Forum Message <> Reply to Message

Actually, I don't get the same result. I used b8 (age of children) and used tab b8 and it showed 10935 observations but when i use svy:tab b8 it shows the same number of observations 10935 but population size is 11040. Whereas, when i use b4 (gender of child) and use tab b4, it shows 11763 observations and when I use tab b4 if b5==1, then it shows 10935, same as b8. But, here is the interesting part, when I use svy: tab b4 it shows observations as 11763 and population size as 11978 but when I use svy: tab b4 if b5==1, then I get 11040 in population size. So, I think svy has a role to play here. But, I don't know what is 11040 then? If it is the count representing children under 5 years of age and alive then why didn't tab b4 if b5==1 give the same number?

Subject: Re: svy command and descriptives
Posted by Bridgette-DHS on Tue, 11 Oct 2016 10:50:36 GMT
View Forum Message <> Reply to Message

Another response from Tom Pullum:

Quote:You are talking about the difference between weighted and unweighted frequencies. There is a fundamental misunderstanding of the weights here. When you do not use weights, you get the unweighted number of cases in the data. When you do use weights, you get the weighted number of cases in the data. They do not match. The weights have been constructed in such a way that in the IR file (PKIR61FL.dta) the total number of weighted and unweighted cases will be the same. You can check this with the IR file by comparing "tab v007" and "tab v007[iweight=v005/1000000]". Both of them will give a total of 13,558 women. The mother's weights (v005) have been transferred directly onto the KR file. Within the KR file, the corresponding totals will NOT match. If in the KR file you enter "tab v007" you will get an unweighted total of 11,763 children. If you enter "tab v007[iweight=v005/1000000]" you will get a weighted total of 11,977.381 children. These two numbers are different, simply because the number of children born in the past five years is not statistically independent of the weight variable. The difference is very small, but some difference is virtually inevitable. The weighted number of cases is not a "population size". Proportions and means calculated with weights will be better estimates of population characteristics than the UNweighted proportions and means would be, but you cannot interpret a weighted frequency as a "population size."

Subject: Re: svy command and descriptives
Posted by shayankhan on Tue, 11 Oct 2016 11:18:39 GMT
View Forum Message <> Reply to Message

Ah okay. I got it now. So, I have to use the weighted data in the KR file if I want the correct representation of the population. I was just confused because when we use svy commands, the output shows the weighted number of cases as population size. Thanks again.