Subject: Weighting data
Posted by pjoshi on Tue, 30 Sep 2014 10:57:21 GMT
View Forum Message <> Reply to Message

Hi,

I am working with the 2011 data set for Nepal and I had a question regarding the sample weights. I have been performing very basic calculations to try and replicate the results of DHS 2011 report to understand the data. However, my numbers don't match the numbers on the report and I want to check if I am missing something important. In order to apply the sample weights when performing calculations on STATA, is the following process correct?
Example: to calculate the mean number of household members

 gen weight= HV005/1000000

svyset HV021 [pweight=weight], strata(HV023)

I get the following message after this:

```
    pweight: weight
        VCE: linearized
  Single unit: missing
     Strata 1: HV023
       SU 1: HV021
      FPC 1: <zero>
```
Next, I entered
svy: mean V009

I get a mean of 4.6 while the number of the DHS report is 4.4. I would really appreciate any help on this.

Thanks!

Subject: Re: Weighting data
Posted by Trevor-DHS on Fri, 03 Oct 2014 15:34:36 GMT
View Forum Message <> Reply to Message

I had to guess at what you are trying to do. I presume that you are working with the household recode (HR) dataset (NPHR60FL.dta) and you are trying to match the number of household members. Your message doesn't actually state what you are doing, but I think this is right.

In your commands below the variable names should be lower case, but I assume that you used

these correctly as you would have received error messages from stata, and that the mistake was just in retyping them here.
I am also guessing the you meant to use hv009 in the svy: mean command and not V009.

hv009 gives the total number of members listed in the household roster, including both usual members of the household and visitors that slept in the household the previous night.
To match the numbers in the report (I'm guessing you are referring to table 2.10, but you didn't state which numbers and which indicators you were trying to reproduce) you would need to use only the de jure (usual) members of the household. The count of de jure (usual) members of the household is given in hv012. Thus use:
svy: mean hv012
to produce the result in table 2.10

Subject: Re: Weighting data
Posted by Bridgette-DHS on Mon, 06 Oct 2014 14:23:51 GMT
View Forum Message <> Reply to Message

Here is additional comments from Senior Specialist, Tom Pullum:

If all you want to do is to calculate mean household size, then you do not need to make an adjustment for clusters and strata. That adjustment only affects the estimates of standard errors. You just need the adjustment for weights. You can use this with the HR file (NPHR60FL.dta):

summarize hv009 [iweight=hv005/1000000]

Or you can do the following (for pweights, it is not necessary to divide by 1000000):

svyset [pweight=hv005]
svy: mean hv009

Both of those will give 4.6289, which is probably what you were getting and does not match with the value in the report, 4.4. The reason you are not getting 4.4 is that the DHS figure is limited to de facto members of households, i.e. respondents for whom hv103=1 rather than 0. Try the following, with the PR file rather than the HR file, i.e. with NPPR60FL.dta:

collapse (sum) hv103 (first) hv005, by(hhid)
summarize hv103 [iweight=hv005/1000000]

This will give a mean of 4.39408, which matches the report.

Subject: Re: Weighting data
Posted by Trevor-DHS on Mon, 06 Oct 2014 15:47:15 GMT
View Forum Message <> Reply to Message

Just a further clarification in case the two responses are confusing. Tom's response is a way of getting the mean number of de facto household members. To get the number of de jure household members (as used in the table in the report) you would use hv102 (usual member) instead of hv103 (slept in the household the previous night) in Tom's response.

Similarly, you can just use the HR file and get the mean of hv012 for mean number of de jure household members, or use hv013 for the mean number of de facto household members.
Mean number of de jure household members: 4.44513
Mean number of de facto household members: 4.39408

Table 2.10 in the survey report provides the mean number of de jure household members.