Subject: Weighting Couples HIV data
Posted by DaniD on Mon, 19 Oct 2015 16:10:37 GMT
View Forum Message <> Reply to Message

Hi!

I am working with the couples data merged with HIV data. I am wondering which HIV weight (men's or women's) I should be using when I am running analyses that involve both men and women.

Thanks!
Dani

Subject: Re: Weighting Couples HIV data
Posted by Bridgette-DHS on Wed, 21 Oct 2015 17:05:46 GMT
View Forum Message <> Reply to Message

Following is a response from Senior DHS Data Processing Specialist, Noureddine Abderrahim:

There is no rule of thumb to do this; it all depends on what you want to analyze. In a case where you need to analyze the data using both the men and the women HIV status, you will have to define a new variable in which you will include the status of the couple:

Both are HIV negative
Both are HIV positive
One of them is HIV positive
The last category could be defined as woman HIV negative and men HIV positive and woman HIV positive and man HIV negative.

Subject: Re: Weighting Couples HIV data
Posted by DaniD on Mon, 02 Nov 2015 20:52:18 GMT
View Forum Message <> Reply to Message

Thank you for your reply. I am still a bit confused. I have the HIV couple status set up as you described it for my analyses but are you saying that I should be creating different weights based on the couple's HIV status?

Thank you for any and all help with this.

Dani

Subject: Re: Weighting Couples HIV data

Hi again,

Actually this might be a good question for Tom Pullum since he has actually done this exact type of weighting.

Here is how I am thinking about the weighting issue and maybe he or someone else can confirm that this is correct.

So I am doing an analysis of HIV-tested couples in xx African countries. I am planning to use the WOMEN'S HIV weights, since they were the ones selected for the DHS samples. I will then follow the strategy for combining countries that Tom describes in another topic thread which is copied below:

Following is a response from Senior DHS Stata Specialist, Tom Pullum
Any analysis using hiv03 (result of the HIV test) should be weighted with hiv05. Once you have merged with the AR file, you should ignore v005 or mv005 or hv005.

Your within-survey analyses are fine with the original hiv05 as the weight. Any adjustment to the weights related to pooling would be by a survey-specific multiplier and would have no effect on within-survey estimates.

My preferred way to handle the renumbering of clusters and strata in a pooled file is to use the "egen group" command. Within each survey, the cluster variable is always v001 (which is duplicated as v021). The stratum variable does not always have the same number and it is not always even named correctly. The strata are virtually always the combinations of region x v025 (v025 is urban/rural). I would find or construct that variable and then rename it as "strata", e.g. "gen strata=v022". You also need a unique identifier for "survey". You cannot rely on v000 for this, because v000 is a 3-character string such as "NG5", where "NG" is the country id and "5" is the phase of DHS. Sometimes there will be two surveys in the same phase, and v000 will be the same for both of them. (This is not an issue if you are using just one survey per country.) Anyway, you will need a line such as

egen cluster_pooled=group(survey v001)
egen strata_pooled=group(survey strata) and then you will have the unique identifiers.

To give equal weight to each survey, you need lines such as these FOR EACH SURVEY SEPARATELY:

```
scalar TOTWT=1000000
quietly summarize hiv05
scalar T=r(sum)
gen hiv05r=hiv05*TOTWT/T
```

You can do this adjustment before the pooling, or put those lines in a loop after the pooling, but just be sure that the recoding is survey-specific. These lines will remove the arbitrary factor of 1000000 from the original hiv05 and will give each survey an arbitrary TOTAL weight of 1000000.

(That number could be anything you want.) This approach will give the same weight to every survey, regardless of the population of the country or the size of the sample. You need to make it very clear that your regional estimates were calculated that way. If, say, you wanted to weight each survey in proportion to its population size, you would replace "TOTWT" with the country's total population or the population age 15-49 or something like that.

If you think there are problems with this strategy and/or can propose a better one, please let me know.

Thanks so much!

Dani

Subject: Re: Weighting Couples HIV data
Posted by Bridgette-DHS on Thu, 12 Nov 2015 14:44:37 GMT
View Forum Message <> Reply to Message

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

If you are using couples with the HIV data attached, DHS would recommend that you use hiv05 from the men as the weight variable. We always have worse nonresponse for men than for women, especially for HIV tests. This is contrary to your suggestion. Sorry about that.

Just "off the top of my head" I would suggest that in the couples file you take the log of hiv05 for the women and hiv05 for the men. I would predict that in most surveys the standard deviation of log(hiv05) will be greater for men than for women. That's a way of quantifying the variability in nonresponse.

Stan Becker at Johns Hopkins University has proposed a way to synthesize the weights from men and women in a couples file. Here's a reference: http://uaps2015.princeton.edu/abstracts/151669.