Subject: Sample weight and use of svy command in regression
Posted by rkchettri on Tue, 19 Dec 2017 12:40:19 GMT
View Forum Message <> Reply to Message

Hi DHS experts,
I am using Nepal DHS 2016 data, I am using IR file. For weighting purpose I am using following command:

gen rweight=v005/1000000
svyset v021 [pweight=rweight], strata(v023) vce(linearized) singleunit(missing)

And for the regression purpose I using this command:
svy: logit outcome variable (eg 4anc) predictor variable (eg age of women) , or

I am want to be sure either I am doing correct or not?

Regards,
Resham

Subject: Re: Sample weight and use of svy command in regression
Posted by Bridgette-DHS on Thu, 21 Dec 2017 22:55:30 GMT
View Forum Message <> Reply to Message

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

Yes, your command is fine. You could simplify it slightly in two ways and get the same results. First, for pweights it is not necessary to divide v005 by 1000000. Stata will do this automatically, because pweights are automatically re-scaled to have a mean of 1. Second, vce(linearized) is a default and does not need to be specified. Thus your svyset could be simply this: svyset v021 [pweight=v005], strata(v023) singleunit(missing).

Subject: Re: Sample weight and use of svy command in regression
Posted by rkchettri on Fri, 22 Dec 2017 00:25:07 GMT
View Forum Message <> Reply to Message

Thank you for response and new ideas.
I checked and found the same results in regression analysis. But while run frequencies without dividing v005 by 1000000, I found different results, for example,

. tab m14ANCcat[iweight=v005]

RECODE of m14_1 (number of antenatal visits during pregnancy)	Freq.	Percent	Cum.
No ANC	71978871	3.64	3.64
1-3ANC	504713483	25.52	29.15
4ANC	1401330967	70.85	100.00
Total	1978023321	100.00	

This is different when v005 is divided by 1000000

. tab m14ANCcat[iw=v005/1000000]

RECODE of m14_1 (number of antenatal visits during pregnancy)	Freq.	Percent	Cum.
No ANC	71.978871	3.64	3.64
1-3ANC	504.713483	25.52	29.15
4ANC	1,401.331	70.85	100.00
Total	1,978.0233	100.00	

So, to calculate the frequencies, we need to v005 by 1000000. Am I right?

Looking forward to your response.

With kind regards,
Resham

Subject: Re: Sample weight and use of svy command in regression

Posted by Bridgette-DHS on Fri, 22 Dec 2017 12:31:44 GMT
View Forum Message <> Reply to Message

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

Yes, that's right. When using iweight you do need to divide by 1000000. It is only with pweight that Stata will automatically re-scale the weights to have a mean of 1.

Subject: Re: Sample weight and use of svy command in regression
Posted by rkchettri on Wed, 10 Jan 2018 12:55:20 GMT
View Forum Message <> Reply to Message

Dear DHS expert team,
One more thing to ask regarding the svyset command and final regression model. Say if we set the svy command using: svyset v021 [pweight=v005], strata(v024) singleunit(missing). Can I include strata (v024) (in my case v023 is provinces which has important predictor) as explanatory in the final regression model?
I have conducted multivariate logistic regression analysis using command like this.
svy: logit outcome variable (eg 4ANC) varlist of explanatory variables (eg i.ethnicicity i.wealth rank .....v024), or

svy: logit 4anc i.v024 i.v025 i.v130 i.v131,or

Is this right command ?

Waiting for advice.

Best

Subject: Re: Sample weight and use of svy command in regression
Posted by Bridgette-DHS on Thu, 11 Jan 2018 13:16:27 GMT
View Forum Message <> Reply to Message

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

Yes, a variable that is in the svyset command can also be used in the analysis. There is no problem with your logit regression command. It's likely that religion and ethnicity are associated with place of residence and they may have different effects in different areas. But those are analytical issues.